



n°30
Février
2015

La lettre IN2P3 Informatique

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

A LA
UNE

Smilei, un code PIC performant et open source



De nombreux laboratoires du Plateau de Saclay se sont fédérés autour du projet CILEX (Centre Interdisciplinaire de la Lumière Extrême) qui hébergera le futur laser APOLLON. L'objectif est d'étudier la matière sous forme de plasma lorsqu'elle est irradiée par des lasers d'intensité extrême encore inédite. Afin d'accompagner ce très grand effort expérimental, il a été décidé de développer Smilei, un nouveau code de simulation pour lequel les physiciens bénéficient de la collaboration de nombreux acteurs experts en calcul scientifique. Et la

version 1.0 vient tout juste de sortir.

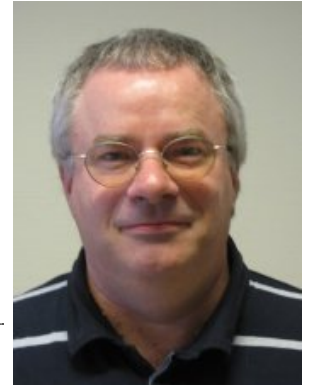
La spécificité première de Smilei est qu'il est le produit d'une collaboration entre différents laboratoires, mais surtout, qu'il associe des équipes ayant des expertises très différentes et complémentaires. Le petit groupe initial de physiciens et d'experts en calcul intensif du LLR, du LULI et de la Maison de la Simulation a, en effet, été progressivement renforcé par l'IDRIS, le LRI et le LPP.

Ce projet regroupe finalement des physiciens des plasmas, des astrophysiciens, des numériciens et des spécialistes du calcul intensif. Ainsi, l'ambition de l'équipe SMILEI dépasse le seul cadre du projet CILEX. Il s'agit de développer un code permettant la simulation d'une large gamme de phénomènes physiques et qui soit capable (...)

[lire la suite](#)

Interview

"Nous ne pouvons pas nous permettre de ne pas construire des passerelles vers les mondes du HPC et du Cloud"



Laurent Duflot, responsable scientifique de LCG-France

[lire l'interview](#)

■ Événement



Journée des Expériences au CC-IN2P3 : le cycle de vie des données en question

Rendez-vous de début d'année, la Journée des Expériences a eu lieu le 22 janvier au Centre de Calcul de l'IN2P3 au lendemain de la réunion d'arbitrage sur les demandes exprimées pour 2015 par les groupes. Pierre-Etienne Macchi, directeur du CC-IN2P3, a d'ailleurs pu annoncer que le budget restait constant par rapport à 2014 et que les demandes de ressources effectuées devraient être couvertes.

Moment de rencontre entre les équipes du Centre de Calcul et la communauté des utilisateurs, nous avons souhaité cette année une plus grande interaction avec les personnes présentes en proposant deux sessions plus dédiées à la discussion. La première avait pour thème le cycle de vie des données. Animée par Jean-Yves Nief, responsable de l'équipe Stockage, elle pose le problème du devenir des données pour les utilisateurs/groupes inactifs

■ Bio-informatique



eTRIKS - Une plateforme de Bio-Informatique au CC-IN2P3

eTRIKS [1] est un partenariat public-privé européen établi dans le cadre de l'initiative IMI [2] visant à accélérer le développement de nouveaux médicaments plus performants et plus sûrs pour les patients. Elle vise spécifiquement trois domaines : • Réduire les coûts de calcul et d'exploitation des données et des bases de connaissances de recherche translationnelle [3] ; • Permettre aux non-statisticiens d'effectuer des analyses exploratoires ; • Faciliter l'analyse croisée d'études venant de domaines différents, par exemple de cohortes de patients avec des données d'essais cliniques. Plus spécifiquement dans le cas d'eTRIKS, le but du projet quinquennal est la gestion et le développement d'une plateforme collaborative d'analyse bio-informatique de données cliniques accessible via Internet. Plusieurs projets nationaux ou européens ont rejoint eTRIKS pour bénéficier du déploiement de cette nouvelle plateforme. C'est le cas, entre autres, de Abrisisk, Oncotrack, U-Biopred ou encore récemment de Bioaster. Partenaire clé, le

■ Collaboratif



Gitlab au CC-IN2P3 : la parole est au social coding

Le CC-IN2P3 propose depuis plusieurs années différents outils de gestion de code. Qu'ils soient adeptes de CVS, SVN, Mercurial ou Git, près de 1200 personnes utilisent aujourd'hui ces logiciels avec lesquels près de 300 dépôts de code sont gérés sur scm.in2p3.fr (aussi connu sous les noms svn.in2p3.fr et git.in2p3.fr). Mais la mode est aujourd'hui au social coding, c'est-à-dire au développement collaboratif de code avec les outils adéquats pour partager, commiter et travailler ensemble en un seul clic. Dans ce paysage d'outils de gestion de code, un petit nouveau vient donc de faire son apparition : Gitlab, un clone du célèbre GitHub. Outre l'hébergement de dépôts Git, il propose également des outils annexes pour la gestion de projets, tels qu'un gestionnaire de bugs ou un wiki pour la documentation. Mis en place au CC-IN2P3 depuis le mois de mai et en production au 1er janvier 2015, Gitlab a déjà ses aficionados puisque 108 utilisateurs sont déjà répertoriés. Parmi les logiciels

■ Formation



ENVOL 2014 - Méthodes de test et validation des logiciels

ENVOL est une Action Nationale de Formation (ANF) organisée par le CNRS (via PLUME, DevLOG, la DSI et des laboratoires du CNRS) proposant des stages de formation (des écoles) tous les deux ans. Le développement logiciel est un processus complexe allant bien au delà de l'écriture de code logiciel. Les écoles ENVOL proposent chaque année une thématique autour de ces développements pour la recherche scientifique. Le public visé par ENVOL est large : ingénieurs et chercheurs (y compris doctorants) ayant une activité de développement logiciel. Les présentations des différentes écoles sont disponibles sur le site PLUME. L'édition 2014, qui s'est déroulée du 18 au 21 novembre 2014 à La-Londeles-Maures, proposait une formation aux méthodes de test et validation des logiciels. Cette édition était organisée en collaboration avec l'INRA, VetAgro-Sup et Inria. Cette thématique a soulevé un intérêt certain puisqu'elle a attiré 53 participants de

Lu pour vous

BIOASTER démultiplie ses capacités bioinformatiques en signant un partenariat stratégique avec le CC-IN2P3
Le stockage massif de données et la puissance de calcul de la physique subatomique et (...)
[en savoir plus](#)

A noter

Atrium vous attend !
La nouvelle plateforme documentaire de l'IN2P3 est prête à accueillir vos projets, dans un esprit (...)
[en savoir plus](#)

Agenda

8 au 11 déc - 11èmes Journées RESeaux (JRES)
Les 11èmes Journées RESeaux (JRES) auront lieu du 8 au 11 décembre 2015 à Montpellier. Les JRES (...)
[en savoir plus](#)

Archives

Abonnement

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

Proposer un article

Vous souhaitez proposer un article ? Envoyez un mail à IN2P3informatique@in2p3.fr.



(ou en fin d'activité). Que doit-on faire de toutes ces informations stockées sur les différents supports proposés par le Centre ? Faut-il les conserver et pour combien de temps (...)

[lire la suite](#)

Centre de Calcul de l'IN2P3 contribue au projet de (...)

[lire la suite](#)

hébergés sur Gitlab, citons par exemple les modules Puppet qui gèrent la configuration (...)

[lire la suite](#)

diverses disciplines, dont 4 de l'IN2P3, ainsi que l'organisateur. Le programme était composé de (...)

[lire la suite](#)



n°30
Février
2015

La lettre IN2P3 Informatique

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

A LA
UNE

Smilei, un code PIC performant et open source



De nombreux laboratoires du Plateau de Saclay se sont fédérés autour du projet CILEX (Centre Interdisciplinaire de la Lumière Extrême) qui hébergera le futur laser APOLLON. L'objectif est d'étudier la matière sous forme de plasma lorsqu'elle est irradiée par des lasers d'intensité extrême encore inédite. Afin d'accompagner ce très grand effort expérimental, il a été décidé de développer Smilei, un nouveau code de simulation pour lequel les physiciens bénéficient de la collaboration de nombreux acteurs experts en calcul scientifique. Et la version 1.0 vient tout juste de sortir.

La spécificité première de Smilei est qu'il est le produit d'une collaboration entre différents laboratoires, mais surtout, qu'il associe des équipes ayant des expertises très différentes et complémentaires. Le petit groupe initial de physiciens et d'experts en calcul intensif du LLR, du LULI et de la Maison de la Simulation a, en effet, été progressivement renforcé par l'IDRIS, le LRI et le LPP.

Ce projet regroupe finalement des physiciens des plasmas, des astrophysiciens, des numériciens et des spécialistes du calcul intensif. Ainsi, l'ambition de l'équipe SMILEI dépasse le seul cadre du projet CILEX. Il s'agit de développer un code permettant la simulation d'une large gamme de phénomènes physiques et qui soit capable de tirer profit au maximum des ressources de calcul intensif modernes.

Les travaux de développement du code ont démarré en janvier 2013, et il existe aujourd'hui une version permettant la simulation à une et deux dimensions de divers phénomènes issus de l'astrophysique et de la physique de l'interaction laser-plasma.

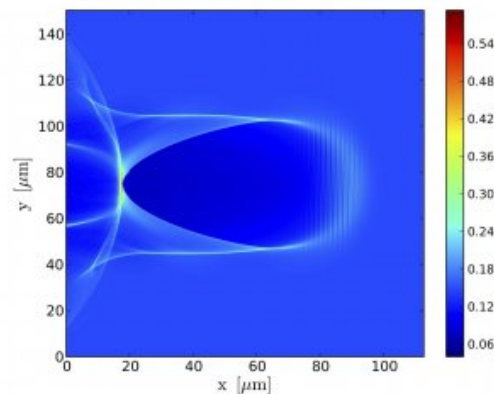
Smilei est un code Particle-In-Cell (PIC) écrit en C++. Il utilise des schémas aux différences finies pour résoudre les équations de Maxwell sur une grille cartésienne régulière. Des objets ponctuels, appelés macro-particules, servent à échantillonner l'espace des phases du plasma et se déplacent librement dans l'ensemble du domaine discrétisé par la grille de calcul en fonction des champs évalués dans leur voisinage. Une fois le déplacement des macro-particules effectué, on peut projeter sur la grille les densités de charge et de courant qu'elles portent puis résoudre à nouveau les équations de Maxwell avec ces nouveaux termes sources.

- La parallélisation est assurée dans un premier temps par une décomposition MPI classique du domaine de calcul. Des communications entre voisins sont nécessaires mais comme elles restent toujours locales, le passage à l'échelle ne pose aucun problème.

- Une parallélisation plus fine au niveau de chaque sous domaine MPI est assurée avec openMP. Les sous domaines sont divisés en paquets de cellules, appelés « clusters », et un tri des particules est effectué pour s'assurer que toutes les particules associées à un même cluster soient contiguës en mémoire. Les clusters peuvent alors être traités indépendamment les uns des autres, et en parallèle, par les threads openMP. Cette parallélisation fine en mémoire partagée est une véritable aubaine pour les cas qui présentent un fort déséquilibre de charge. Celui-ci est dynamiquement réparti entre les threads et son impact sur le temps de calcul global est divisé d'autant. C'est de plus un excellent moyen de faire face à la

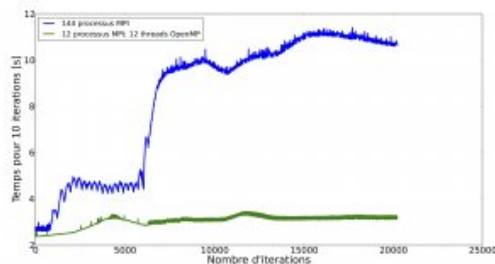
en France.

Ci-dessous figure un exemple de simulation d'accélération d'électrons par sillage laser. La figure nous montre la densité électronique à un instant donné. Cette densité correspond également à la charge locale de calcul car celle-ci est fonction de la densité de macro-particules.



Simulation Smilei. Densité électronique dans le sillage d'un laser qui se propage dans un plasma selon l'axe x © LLR / CNRS

On note la formation d'un « hot spot » à l'arrière (à gauche sur la figure) de la structure dite de « bulle ». Ce point de très forte densité concentre un très grand nombre de macro particules sur seulement quelques cellules et provoque un très fort déséquilibre de charge. La courbe de performance nous montre le temps d'exécution de 10 itérations au fur et à mesure de la simulation.



Temps d'exécution de 10 itérations au fur et à mesure de la simulation pour deux schémas de parallélisation © LLR / CNRS

On voit alors clairement l'apport de la parallélisation openMP. Après 6000 itérations, au moment de la formation de la bulle, quand le hot spot apparaît, les performances en MPI pur, déjà dégradées, finissent de s'effondrer complètement alors que l'on garde de très bons résultats grâce au partage dynamique des tâches entre threads openMP.

L'attractivité de Smilei vient aussi du fait qu'il s'agit d'un projet Open-Source. C'est le ciment qui a permis de regrouper des acteurs venant d'horizons aussi différents. Le code est actuellement disponible sur le [gitLab du LLR](#) dans sa version 1.0 sortie en Novembre 2014. N'hésitez donc pas à aller le tester, à donner un retour aux développeurs, et, pourquoi pas, rejoignez vous même le groupe des « Smilers » en tant que développeur ou utilisateur régulier.

Contact :

tendance actuelle des constructeurs qui introduisent de plus en plus de cœurs dans les processeurs et remettent ainsi au goût du jour les architectures à mémoire partagée.

■ beck@llr.in2p3.fr

■ julien.derouillat@cea.fr

SMILEI a déjà fait ses preuves dans la cour des grands. Le code a bénéficié de 3 millions d'heures de calcul sur le tout nouveau calculateur OCCIGEN du CINES dans le cadre du grand challenge orchestré par GENCI pour l'arrivée de ce nouveau système qui occupe désormais la place de numéro 1

■ mickael.grech@polytechnique.edu

Arnaud BECK (LLR)



n°30
Février
2015

La lettre IN2P3 Informatique

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



"Nous ne pouvons pas nous permettre de ne pas construire des passerelles vers les mondes du HPC et du Cloud"

Laurent Dufлот, responsable scientifique de LCG-France



- Pouvez-vous rappeler brièvement votre parcours ?

Je suis un physicien en physique des particules à l'IN2P3 depuis 1993 et j'ai participé à trois générations de "grosses expériences" de la communauté (ALEPH, DØ et ATLAS) sur la physique de précision du Modèle Standard puis les recherches de nouvelles particules. J'ai également travaillé sur les aspects software et computing de notre discipline. Avant ma thèse, j'ai travaillé un an sur le portage de GEANT (le logiciel de simulation de détecteur) sur une machine parallèle. Il y a une dizaine d'années, j'ai été un des beta-testers de la grille de données de DØ appelée SAM-GRID qui a ensuite été utilisée dans la production d'événements simulés pour l'expérience, surtout sur l'infrastructure de grille américaine OSG mais aussi en Europe avec des outils LCG (LHC Computing Grid). J'ai poursuivi en explorant l'analyse DØ sur la grille LCG. Ces expérimentations m'ont permis à mon arrivée dans ATLAS de porter un logiciel gourmand en temps de calcul sur la grille LCG (à l'époque il n'y avait pas d'outils faciles d'accès pour les utilisateurs). J'ai donc été tout au long de ma carrière sensible à l'importance du software et du computing pour obtenir des résultats de physique dans notre discipline.

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

Je suis le coordinateur scientifique de LCG-France depuis le premier janvier et à ce titre je représente la France dans certaines instances de la collaboration WLCG qui regroupe des représentants des sites de calcul pour les quatre expériences du LHC et des représentants des expériences. Un des objectifs concrets de LCG-France est de maintenir la contribution française au calcul pour le LHC à une part de 8 à 10 % du calcul mondial. Il faut pour cela adapter au mieux nos moyens aux besoins exprimés par les expériences et donc en particulier coordonner les évolutions des sites français : Tier1/Tier2 au CC-IN2P3 et Tier2/Tier3 dans nos laboratoires. Pour cela, je collabore étroitement avec la coordinatrice technique Catherine

Malek depuis la création de LCG-France. Fairouz a une profonde connaissance du projet et de son histoire et prendre sa succession est un challenge !

- Le LHC devrait être remis en production au printemps. Qu'est-ce qui va changer dans cette nouvelle phase de prise de données ?

Le LHC va redémarrer à plus haute énergie, 13 TeV au lieu de 8 TeV. Cela permettra par exemple de produire beaucoup plus de bosons de Higgs et donc de passer d'un mode « découverte et premières explorations de ses propriétés » à un mode d'études plus approfondies. Ce nouveau domaine d'énergie ouvre aussi la voie à des découvertes de nouvelles particules qui auraient été auparavant inaccessibles. Le démarrage sera donc un moment de grande excitation mais aussi de stress car tout doit marcher très vite, du détecteur à l'analyse de données finale. La pression est donc aussi mise sur le calcul : simulation et reconstruction des données, distribution à travers la grille, cycles d'analyse rapide. D'ores et déjà, les productions de simulation des expériences dans les configurations nominales montent en puissance.

- Quels sont les prochains défis à relever en matière d'informatique pour le LHC ?

Lors de l'arrêt prolongé de la machine, les expériences ont travaillé sur leur software (optimisation et adaptation au nouveau détecteur de la simulation et de la reconstruction, format d'analyse de données, etc.) mais aussi sur leur modèle de calcul. La hiérarchie assez rigide Tier1/2/3 imaginée avant le démarrage du LHC pour assurer un calcul « bien orchestré » laisse la place à un modèle plus dynamique. Si le rôle des Tier1 reste central, les meilleurs Tier2 doivent démontrer une grande disponibilité et une interconnexion avec les Tiers2 et Tiers1 « voisins » nettement plus performante que dans le modèle initial. Le placement des données sur la grille devient plus dynamique et basé sur la « popularité » des données. Ces évolutions doivent se poursuivre pendant la prise de données tout en garantissant un système suffisamment robuste pour résister au « rush » sur les données. Si en 2015, la quantité de données prévue est comparable à celle prise lors du Run1 du LHC, le volume de données va vite croître et il faut être prêt. Dans le même temps, WLCG étudie comment simplifier les opérations de la grille et des sites : diminution du nombre de services ou des protocoles à supporter, convergence des expériences vers des solutions communes, installation et support plus faciles, etc.

L'informatique du LHC doit aussi se placer dans le contexte global du calcul scientifique dans le monde. Le calcul sur super-calculateur haute performance (HPC) est un gros acteur du calcul scientifique, en général développé pour d'autres problématiques que les nôtres (machines très puissantes avec beaucoup de mémoire mais pas forcément construites pour des échanges de données aussi importants que les nôtres). Le « cloud computing » s'est énormément développé dans l'industrie et prend son envol dans le domaine du calcul scientifique. Même si la grille dédiée reste au coeur du calcul pour le LHC, nos besoins sont tels que nous ne pouvons pas nous permettre de ne pas construire des passerelles vers les mondes du HPC et du Cloud.

- Enfin, le Boson de Higgs découvert, quelles sont les prochaines découvertes pressenties dans le cadre des expériences du LHC ?

La montée en énergie offre la possibilité de produire des particules trop massives pour être produites jusqu'ici. Il existe donc de réelles possibilités de découverte si la nature nous fait ce plaisir ! Je pense en particulier aux particules prédites dans le

Biscarat. Nous avons la chance d'avoir collaboré ensemble dans l'expérience D0 lorsque nous étions tous deux sur le site de Fermilab, elle travaillant sur la simulation et moi comme coordinateur du software.

Les opérations au jour le jour de la grille LCG sont très prenantes mais il nous faut prendre le temps de réfléchir aux évolutions futures dans notre domaine. Catherine et moi-même encourageons la participation des français dans les groupes de travail sur les outils de la grille, aux décisions stratégiques sur les évolutions majeures et à la veille technologique. Il nous faut également nous insérer dans le contexte du calcul scientifique en France et en particulier des nouvelles politiques de calcul du ministère et du CNRS. J'ai repris le flambeau de la coordination scientifique qui a été tenu en main par Fairouz

cadre des théories supersymétriques. Ces théories sont séduisantes car elles permettent de combler quelques failles du Modèle Standard de la physique des particules et de proposer une particule candidate pour la matière noire de l'univers. Pour colmater les brèches de manière théoriquement satisfaisantes, certaines des particules supersymétriques devraient avoir une masse inférieure à quelques TeV et donc seraient dans le domaine accessible au LHC. La découverte de particules supersymétriques serait une avancée majeure pour la discipline !

PROPOS RECUEILLIS PAR GAELLE SHIFRIN



n°30
Février
2015

La lettre IN2P3 Informatique

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



Journée des Expériences au CC-IN2P3 : le cycle de vie des données en question



© CC-IN2P3 / CNRS

Rendez-vous de début d'année, la Journée des Expériences a eu lieu le 22 janvier au Centre de Calcul de l'IN2P3 au lendemain de la réunion d'arbitrage sur les demandes exprimées pour 2015 par les groupes. Pierre-Etienne Macchi, directeur du CC-IN2P3, a d'ailleurs pu annoncer que le budget restait constant par rapport à 2014 et que les demandes de ressources effectuées devraient être couvertes.

Moment de rencontre entre les équipes du Centre de Calcul et la communauté des utilisateurs, nous avons souhaité cette année une plus grande interaction avec les personnes présentes en proposant deux sessions plus dédiées à la discussion.

La première avait pour thème le cycle de vie des données. Animée par Jean-Yves Nief, responsable de l'équipe Stockage, elle pose le problème du devenir des données pour les utilisateurs/groupes inactifs (ou en fin d'activité). Que doit-on faire de toutes ces informations stockées sur les différents supports proposés par le Centre ? Faut-il les conserver et pour combien de temps ? Ce problème s'adresse aussi à l'ensemble des groupes actifs car cette gestion doit être prise

en compte le plus tôt possible afin de ne pas être pris au dépourvu. La discussion a permis de poser les enjeux, de montrer les bénéfices que pourrait apporter un plan de gestion des données et d'avoir un premier retour de la part des utilisateurs. Le sujet est complexe et fera l'objet d'échanges futurs avec la communauté de l'IN2P3.

S'en est suivie une « table ronde » pour laquelle nous avons demandé que les participants à la réunion proposent des sujets. Malgré le peu de propositions reçues, nous avons pu échanger sur le calcul multicœur en particulier, et les évolutions au niveau du stockage. 2015 verra par ailleurs un meilleur support du GPU au Centre de Calcul puisque des demandes officielles ont pour la première fois été émises lors des requêtes de ressources.

La journée s'est terminée par les présentations de deux expériences. Vincent Poireau du LAPP nous a parlé d'AMS (Alpha Magnetic Spectrometer), détecteur de particules en orbite depuis 2011, mesurant le flux de rayons cosmiques pour la recherche de matière noire et d'antimatière. En terme de calcul, un nouveau format de données ACQt est actuellement en cours d'élaboration au sein d'AMS en remplacement de Root et devrait être testé au Centre. Chercheur à l'IPNL, Olivier Stezowski a, pour sa part, présenté l'expérience de physique nucléaire AGATA (Advanced GAMMA-ray Tracking Array), spectromètre gamma pour l'étude du noyau. Collaboration multi-projets et disposant de peu de ressources, elle tire notamment bénéfice des multiples outils collaboratifs mis à disposition par le CC-IN2P3 tout en utilisant les ressources de calcul et stockage. Olivier Stezowski a pointé le besoin, notamment pour ce type de collaboration, de plus d'informations sur les évolutions en cours au Centre et d'un catalogue de service, qui est actuellement en chantier.

Vous pouvez retrouver les présentations de cette journée à l'adresse :

<https://indico.in2p3.fr/event/10875/timetable/#20150122>.

David BOUVET (CC-IN2P3)



n°30
Février
2015

La lettre IN2P3 Informatique

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



eTRIKS - Une plateforme de Bio-Informatique au CC-IN2P3

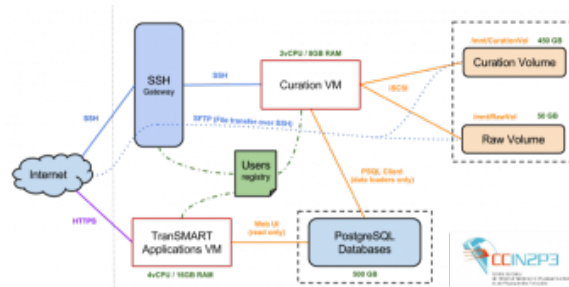


eTRIKS [1] est un partenariat public-privé européen établi dans le cadre de l'initiative IMI [2] visant à accélérer le développement de nouveaux médicaments plus performants et plus sûrs pour les patients. Elle vise spécifiquement trois domaines :

- Réduire les coûts de calcul et d'exploitation des données et des bases de connaissances de recherche translationnelle [3] ;
- Permettre aux non-statisticiens d'effectuer des analyses exploratoires ;
- Faciliter l'analyse croisée d'études venant de domaines différents, par exemple de cohortes de patients avec des données d'essais cliniques.

Plus spécifiquement dans le cas d'eTRIKS, le but du projet quinquennal est la gestion et le développement d'une plateforme collaborative d'analyse bio-informatique de données cliniques accessible via Internet. Plusieurs projets nationaux ou européens ont rejoint eTRIKS pour bénéficier du déploiement de cette nouvelle plateforme. C'est le cas, entre autres, de Abirisk, Oncotrack, U-Biopred ou encore récemment de Bioaster.

Partenaire clé, le Centre de Calcul de l'IN2P3 contribue au projet de par son expertise sur l'hébergement de grands volumes de données et la fourniture de ressources pour leur traitement. La plateforme inclut plusieurs composantes, un serveur public qui est la vitrine démontrant la capacité du projet à analyser des données pour divers types de pathologies ainsi qu'une composante spécifique à chacun des projets collaborant avec eTRIKS. Un module élémentaire de ce déploiement est représenté dans la figure ci-dessous. Il est reproduit pour chaque nouveau projet qui rejoint eTRIKS.



L'architecture de la plateforme est basée sur une architecture de Cloud Openstack sur laquelle les différentes machines virtuelles exécutent des tâches spécifiques. L'accès est accordé au travers d'un mécanisme d'authentification basé sur un catalogue LDAP définissant le rôle et le degré de privilège de chaque utilisateur. L'utilisateur final peut accéder à la lecture et à l'analyse des données à travers d'un logiciel applicatif, **tranSMART**.

Après deux ans de collaboration, le CC-IN2P3 héberge, avec le serveur public, deux projets partenaires en recherche translationnelle : Abirisk (prédiction des risques d'immunisation anti-biopharmaceutique) et Oncotrack (caractérisation des marqueurs biologiques pour la prédiction des tumeurs). Il définit le déploiement et fournit le matériel et l'expertise nécessaire au maintien de la structure cloud décrite ci-dessus et aux tests et mises-à-jour nécessaires pour les outils en jeu.

De plus, la collaboration du Centre de Calcul propose la possibilité d'un support « à distance » pour des projets partenaires d'eTRIKS qui ont choisi un engagement plus indépendant, mais qui veulent bénéficier de l'expertise et des connaissances du partenariat. Le portail opérationnel de eTRIKS, géré à 100% par les membres du CC-IN2P3 engagés dans le projet, est une synthèse concrète de tous les services et du support déployés : <https://portal.etriks.org/Portal/>

Gino MARCHETTI et Ghita RAHAL (CC-IN2P3)

[1] European Translational Information and Knowledge Management Services (Services européens d'information translationnelle et gestion de connaissances)

[2] Innovative Medicines Initiative (Initiative pour des Médicaments Innovants)

[3] La recherche translationnelle est à l'interface entre la recherche fondamentale et clinique. Elle représente une passerelle entre la recherche orientée vers les patients et la recherche à visée cognitive.



n°30
Février
2015

La lettre IN2P3 Informatique

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



Gitlab au CC-IN2P3 : la parole est au social coding



Le CC-IN2P3 propose depuis plusieurs années différents outils de gestion de code. Qu'ils soient adeptes de CVS, SVN, Mercurial ou Git, près de 1200 personnes utilisent aujourd'hui ces logiciels avec lesquels près de 300 dépôts de code sont gérés sur scm.in2p3.fr (aussi connu sous les noms svn.in2p3.fr et git.in2p3.fr).

Mais la mode est aujourd'hui au social coding, c'est-à-dire au développement collaboratif de code avec les outils adéquats pour partager, commiter et travailler ensemble en un seul clic.

Dans ce paysage d'outils de gestion de code, un petit nouveau vient donc de faire son apparition : Gitlab, un clone du célèbre GitHub. Outre l'hébergement de dépôts Git, il propose également des outils annexes pour la gestion de projets, tels qu'un gestionnaire de bugs ou un wiki pour la documentation.

Mis en place au CC-IN2P3 depuis le mois de mai et en production au 1er janvier 2015, Gitlab a déjà ses aficionados puisque 108 utilisateurs sont déjà répertoriés. Parmi les logiciels hébergés sur Gitlab, citons par exemple les modules Puppet qui gèrent la configuration de certains services du Centre.

Petit passage en revue (non exhaustif) des fonctionnalités de Gitlab :

- l'utilisateur peut s'enregistrer, puis, après validation du compte par les administrateurs, immédiatement commencer à créer des dépôts. L'autonomie de l'utilisateur n'en est donc que renforcée.
- l'authentification par une fédération d'identités via le système 'shibboleth' : il est conseillé aux collaborateurs académiques de s'enregistrer et s'authentifier ainsi. Les utilisateurs non-académiques ou non-référencés (ex : visiteurs étrangers)

doivent créer un compte local.

- la gestion des groupes et des droits.
- la possibilité de dupliquer un dépôt ('fork') en vue de le modifier et reverser cette modification sous forme de 'pull request' (ou 'merge request' dans le jargon Gitlab) : il devient donc plus simple de contribuer à un code. Un point essentiel quand on veut travailler en collaboration.
- le système de revue et modification du code en ligne
- la possibilité de s'abonner aux fils de news de surveillance de codes.

Enfin, si vous hésitez entre la Forge (Redmine) ou Gitlab, sachez que Gitlab est plutôt un produit destiné à la gestion de codes, tandis que la Forge apparaît plus adaptée au suivi de projets. Les deux peuvent toutefois s'intégrer, le code étant hébergé dans Gitlab, les tickets et le wiki dans la Forge.

Déployé au CC-IN2P3, Gitlab est donc maintenu, mis à jour tous les mois en horaire décalé (ce qui n'est pas anodin car le produit est en pleine évolution).

Le service Git (git.in2p3.fr) et Gitlab n'offrent pas exactement les mêmes services : Gitlab apporte l'autonomie et la collaboration, Git offre un service d'hébergement plus nu (pas d'interface web de gestion des dépôts), mais aussi l'accompagnement, le conseil et la personnalisation, pour des usages éventuellement plus complexes. Les dépôts peuvent être migrés d'un service à l'autre.

Les expériences, quelle que soit leur taille, devraient être séduites par Gitlab : l'utilisateur est autonome dans la gestion de ses codes, l'utilisation est facilitée grâce à l'interface web, et les fonctionnalités de travail collaboratif sont nombreuses.

Gitlab s'apparente d'ailleurs sur ce point à Gitolite, aux fonctionnalités similaires pour la gestion des droits sur les dépôts, mais avec une gestion en ligne de commandes. Pour les groupes intéressés, Gitolite est disponible sur git.in2p3.fr.

Pour plus d'informations :

- <https://gitlab.in2p3.fr>, <https://scm.in2p3.fr>
- <https://forge.in2p3.fr/projects/git...>

Gaëlle SHIFRIN, Foudil BRETEL et Loïc TORTAY (CC-IN2P3)



n°30
Février
2015

La lettre IN2P3 Informatique

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



ENVOL 2014 - Méthodes de test et validation des logiciels



L'édition 2014 de la formation ENVOL a attiré 53 participants de laboratoires CNRS, INRA, universités françaises, Inserm, Irstea, CEA et des apprentis ingénieurs en formation alternante dans une entreprise toulousaine, dans la résidence Odalys à la Londe les Maures. (Photo CC-BY-NC-ND/3.0 Dirk Hoffmann)

ENVOL est une Action Nationale de Formation (ANF) organisée par le CNRS (via **PLUME**, **DevLOG**, la DSI et des laboratoires du CNRS) proposant des stages de formation (des écoles) tous les deux ans.

Le développement logiciel est un processus complexe allant bien au delà de l'écriture de code logiciel. Les écoles ENVOL proposent chaque année une thématique autour de ces développements pour la recherche scientifique. Le public visé par ENVOL est large : ingénieurs et chercheurs (y compris doctorants) ayant une activité de développement logiciel. Les présentations des différentes écoles sont disponibles sur **le site PLUME**.

L'édition 2014, qui s'est déroulée du 18 au 21 novembre 2014 à La-Londe-les-Maures, proposait une formation aux méthodes de test et validation des logiciels. Cette édition était organisée en collaboration avec l'INRA, VetAgro-Sup et Inria. Cette thématique a soulevé un intérêt certain puisqu'elle a attiré 53 participants de diverses disciplines, dont 4 de l'IN2P3, ainsi que l'organisateur.



Le programme était composé de séances plénières (« cours ») et de TP en groupes individuels. (Photo CC-BY-NC-ND/3.0 Dirk Hoffmann)

Pourquoi vérifier un logiciel ?

Les conséquences d'un dysfonctionnement logiciel sont d'autant plus importantes qu'elles sont détectées tardivement. Elles peuvent varier en terme de temps, de coût financier, d'organisation, mais aussi d'utilisation. Elles peuvent en effet aller jusqu'à invalider de grandes quantités de données.

Autour de ce thème, chaque présentation était suivie de séances de TP, illustrant comment utiliser chaque concept. Ces TP étaient déclinés en plusieurs environnements, chaque participant choisissait celui qui lui était le plus familier (langage

de programmation et système d'exploitation). L'équipe d'ENVOL avait mis une machine virtuelle à disposition afin de proposer un environnement de travail commun, stable, facile d'installation et d'utilisation.

L'objectif d'un développement logiciel est de créer un logiciel à la fois fonctionnel et sûr, c'est-à-dire dont le fonctionnement répond aux attentes. La définition de ces attentes intervient dès le début du projet. Elle entraîne très rapidement des efforts en vue de la validation du logiciel. La conception de tests intervient donc dès le début du projet de développement logiciel pour être utile. Elle ne se limite pas à une phase de test unitaire greffée une fois la programmation terminée.



Une partie des TP a été improvisée en mini-groupes qui s'étaient retrouvés autour de leur langage de préférence. (Photo CC-BY-NC-ND/3.0 Dirk Hoffmann)

La conception et l'intégration de ces tests peuvent se faire par différentes méthodes de génie logiciel. La méthode « Agile », en particulier, a été présentée dans le contexte de la recherche. Elle a été accompagnée de témoignages dans le cadre de besoins qui évoluent au cours du temps, et de leur impact sur l'aspect humain.

Différents types de tests ont été présentés, unitaires et fonctionnels, mais aussi tests de sites web, et automatisation des tests au moyen de plateformes d'intégration continue. En lien avec les plateformes embarquées, des TPs de développement sur Arduino ont aussi été proposés.

Pour finir, la formation a été gratifiée de beau temps dans un endroit magnifique, ce qui a favorisé une ambiance de travail agréable.

Liens utiles :

- Lettre ENVOL 2012 : <http://informatique.in2p3.fr/li/spi...>
- Événement INDICO : <https://indico.in2p3.fr/event/10295/>

Philippe GAURON et Oleg LODYGENSKY (LAL)



n°30
Février
2015

La lettre IN2P3 Informatique

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



8 au 11 déc - 11èmes Journées RESeaux (JRES)



Les 11èmes Journées RESeaux (**JRES**) auront lieu du 8 au 11 décembre 2015 à Montpellier.

Les JRES ont lieu tous les deux ans et s'adressent à l'ensemble de la communauté informatique de l'enseignement et de la recherche : architectes et administrateurs systèmes et réseaux, responsables de projets, directeurs des systèmes d'information, responsables ou correspondants de la sécurité du SI, développeurs et utilisateurs.

Les JRES permettent de montrer et de valoriser l'activité importante de notre communauté, sa contribution à l'intégration des sciences et technologies de l'information et la communication dans nos différentes activités métiers.

Ces journées offrent une occasion majeure de présenter, d'échanger et de discuter des enjeux, des aspects techniques et pratiques des technologies et des services innovants mis en œuvre dans nos établissements ou dans nos laboratoires.

Si vous avez eu l'occasion de réaliser des projets liés à la modernisation, fiabilisation ou sécurisation de l'environnement informatique de nos établissements, vos travaux peuvent intéresser le plus grand nombre.

Venez les partager avec la communauté :

- retour d'expérience, bonnes pratiques ;
- évolution technique ;
- déploiement d'un service innovant, d'une technologie novatrice ;
- résultat de maquettes ou de mises en test de produits ou services ; étude prospective.

Vous êtes invités à soumettre des propositions en vous appuyant sur les thèmes de référence des JRES (liste non exhaustive), selon les formats suivants :

- présentations longues (35 min) ;
- présentations courtes (10 min) ;
- posters ;
- tables rondes.

Les lightning talks, qui ont fait leur arrivée remarquée en 2013 et ont eu beaucoup de succès, seront reconduits en 2015. Ils conserveront un traitement à part avec un mode de soumission dédié, allégé et décalé dans le temps.

Les personnes qui souhaitent soumettre une proposition sont invitées à prendre connaissance des instructions aux auteurs. Les propositions doivent être déposées au plus tard le 15 mars 2015.

La langue de la conférence est le français. Néanmoins, les articles rédigés et/ou présentés en anglais sont également acceptés. Les propositions pourront donc être rédigées dans l'une ou l'autre langue.

Cette année, toutes les sessions programmées dans le grand amphithéâtre seront traduites simultanément en anglais pour les quelques participants anglophones à la conférence.

Un stand sera à disposition des associations dans le domaine du logiciel libre, qu'il s'agisse des logiciels libres en général ou d'un logiciel en particulier. Les associations soumettent une brève présentation de leur association au comité de programme qui effectuera une sélection. Les associations intéressées sont invitées à consulter les instructions aux auteurs pour soumettre leur proposition.

Pour l'équipe JRES 2015 :

Anne Facq et Didier Benza, Présidents du Comité de Programme Sabine Jaume-Rajaonia, Présidente du Comité d'Organisation

presidents@jres.org <https://www.jres.org>



n°30
Février
2015

La lettre IN2P3 Informatique

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



BIOASTER démultiplie ses capacités bioinformatiques en signant un partenariat stratégique avec le CC-IN2P3

Lu pour vous

Le stockage massif de données et la puissance de calcul de la physique subatomique et l'astrophysique sont désormais mis à disposition de la microbiologie de la santé et des maladies infectieuses.

L'Institut technologique en microbiologie BIOASTER et le Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules (CC-IN2P3, une Unité de Service et de Recherche du CNRS) ont signé fin 2014 un accord de partenariat qui permettra de parfaire le stockage et l'analyse des données générées par les programmes de R&D de BIOASTER.

Pour plus d'informations, lire l'intégralité du communiqué de presse ci-dessous :



Atrium vous attend !

A noter

La nouvelle plateforme documentaire de l'IN2P3 est prête à accueillir vos projets, dans un esprit ouvert et collaboratif. Cet outil, mis en place sous l'impulsion de la Direction Adjointe Technique de l'Institut, permet à chacun de stocker et partager ses documents simplement. Des référents locaux et une équipe support dynamique sont à votre disposition pour vous accompagner si nécessaire. Espaces personnels, synchronisation avec votre poste de travail, délégation de droits, Home personnalisable sont autant de nouveautés qui vous apporteront un réel bénéfice. Vos données seront totalement sécurisées par leur hébergement au Centre de Calcul.

Atrium est en prod, Atrium n'attend plus que vous !

<https://atrium.in2p3.fr>