



A LA
UNE

Ça bouge au RI3 !



Ce n'est pas si fréquent, parlons un peu du Réseau des Informaticiens de l'IN2P3 et de l'IRFU, le RI3. Pour certains, cela reste une structure floue, sans impact clairement identifié, dont on admet volontiers ne pas bien connaître les

actions, encore moins les rouages. Par contre, on lui attribue, à juste titre, l'organisation des Journées Informatique et leur succès. Cela tombe bien, les JI reviennent en 2016 (fin septembre début octobre, à Clermont) !

Tout d'abord : avouons-le, cet article est écrit à quatre mains, car Christian, qui voit se profiler une heureuse retraite à l'horizon, passe la main à Frédérique pour l'animation du RI3. Mais soyez rassurés (ou restez inquiets, c'est selon ;-), Christian reste là pour surveiller que tout se passe bien, pour encore un peu moins d'un an maintenant.

CCRI du RI3, kesako ? Loin de reposer sur une seule personne, ou deux, l'animation du réseau se discute et se met en place au sein du comité de coordination du réseau (CCRI). Depuis le printemps dernier, la petite dizaine de personnes qui compose ce comité a pris le parti de se réunir au moins une fois par mois, se voulant une structure plus réactive, à même d'interagir aussi avec le directeur du Centre de Calcul, le chargé de mission à l'informatique et le DAS (directeur adjoint scientifique) chargé du Calcul à (...)

Interview
"il faut se garder de penser que nous sommes à part en matière d'informatique"



lire la suite
Thierry Olivier, chargé de mission formation à l'IN2P3
lire l'interview

■ Infrastructure



CentOS Seven, changement de cap

Le système d'exploitation Scientific Linux (SL) [1] est utilisé depuis 2003 par la communauté HEP [2] et par l'IN2P3 pour ses environnements de calcul. Scientific Linux est l'un des clones de la très populaire distribution commerciale Red Hat Enterprise Linux (RHEL) [3]. SL est né d'une initiative du laboratoire américain Fermilab [4], initiative à laquelle le CERN [5] a apporté une forte contribution. Depuis la version 3 de SL, très instable à ses débuts, à la version 7 en 2015, un long chemin a été parcouru. La qualité du conditionnement du système d'exploitation n'a cessé de s'améliorer pour arriver aujourd'hui à un niveau comparable à bon nombre de distributions commerciales. Toutefois, des interrogations émergent sur le cap à tenir. Si le conditionnement de Scientific Linux n'a plus à rougir face à ses compétiteurs, la politique de fourniture des mises à jour semble encore perfectible. Certains la jugent en retrait par rapport à celle de la distribution CentOS [6], autre

■ Communauté



Dans les coulisses de l'école informatique

Chaque année, l'IN2P3 organise une école informatique (voir l'interview de Thierry Olivier dans ce numéro). La préparation de la dernière en date [1] s'est révélée une expérience très enrichissante pour tous ses acteurs. Le thème était très ambitieux, mais l'équipe rassemblée à cette occasion, emmenée par le tandem Thomas Lallart de l'INRA et Foudil Brétel du CC-IN2P3, a relevé le défi, avec un certain panache, je tiens à le souligner. Bien sûr, l'engagement, les compétences et le talent des intervenants ont été les principaux ingrédients de la réussite de cette petite aventure humaine. Mais je voudrais citer trois outils qui ont rendu possible celle-ci, vécue par une équipe dispersée géographiquement. GitLab, le gestionnaire de dépôt de code partagé par un groupe de développeurs, a permis à chaque membre de l'équipe de connaître dès la

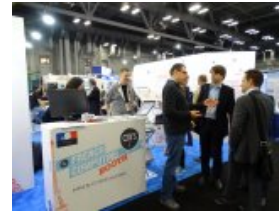
■ Infrastructure



Docker & HENP

Le développement logiciel en physique des hautes énergies – en particulier pour les expériences LHC – met en jeu l'assemblage et l'intégration d'un grand nombre de bibliothèques (externes ou internes, scientifiques ou générales), qui doivent être ensuite compilées, testées et installées sur les machines de production. De plus, et à cause des ressources limitées disponibles, les logiciels des expériences ne sont testés et déployés que sur un jeu très limité de plateformes, architectures, chaînes de compilation et systèmes d'exploitation. Mécaniquement, l'installation d'un environnement de travail sur une machine de développement avec un système légèrement différent (Debian/CentOS/...), ou bien sur une machine de production exploitant une nouvelle architecture, est généralement un processus itératif chronophage, surtout lorsque les performances natives (sans machine virtuelle) sont requises. En effet, les machines virtuelles (VM) ont grandement facilité le

■ Événement



SC'15 : le CC-IN2P3 et l'Ildris représentent le CNRS auprès des grands acteurs du calcul intensif et du big data

Du 16 au 19 novembre, le CC-IN2P3 a présenté pour la 2e fois un stand commun avec l'Ildris à la conférence SuperComputing qui s'est tenue à Austin au Texas (Etats-Unis). Pour la 27e année consécutive, SuperComputing, vitrine majeure du calcul haute performance et du Big Data, a été le rendez-vous incontournable des acteurs industriels et académiques du secteur de l'informatique de pointe pendant une semaine de conférences, de workshops et de présentations. Près de 12 000 visiteurs se sont déplacés pour assister à cet événement qui se veut aussi le plus grand salon au monde où sont présentées les dernières innovations et applications du domaine. Sur le stand du CNRS, était

Remerciements

Deux traits d'union de l'informatique IN2P3-IRFU prennent une retraite bien méritée

Beaucoup d'entre nous l'auront compris ; il s'agit de Pierrick Micout et Joseph Le Foll. Tous les (...)

[en savoir plus](#)

Agenda

Journées LCG France, 14-16 décembre - Villeurbanne (CC-IN2P3)

Les prochaines journées LCG France se dérouleront au Centre de Calcul de l'IN2P3 du lundi 14 au (...)

[en savoir plus](#)

JRES 2015, du 8 au 11 décembre - Montpellier

Les 11èmes Journées Réseau se tiendront au Corum de Montpellier du 8 au 11 décembre. Plus (...)

[en savoir plus](#)

Archives

Abonnement

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

Proposer un article

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



clone populaire de RHEL. Des interrogations sur la pérennité de l'offre Scientific Linux sont également très présentes. Le (...)

[lire la suite](#)

conception, commenter, modifier ou compléter la production de ses collègues. Pas seulement du code, mais aussi et en l'occurrence surtout, des énoncés des exercices et des présentations. Rappelons que le CC-IN2P3 offre un service GitLab national, et (...)

[lire la suite](#)

développement et le déploiement d'applications d'une part, ainsi que la gestion et l'optimisation de l'utilisation de clusters de calcul d'autre part. Cependant, la virtualisation de (...)

[lire la suite](#)

présentée l'infrastructure du CC-IN2P3 et de l'Idris, en mettant l'accent sur la complémentarité de leur offre HPC et HTC (pour High Troughput Computing). Un focus sur le traitement des données pour le LSST était également là pour rappeler le rôle primordial joué par le (...)

[lire la suite](#)



Equipe

Directeur de la publication : Alain FUCHS.

Responsables éditoriaux : Giovanni LAMANNA et Pierre-Etienne MACCHI.

Comité de rédaction : Frédérique CHOLLET LE FLOUR, Virginie DELEBARRE DUTRUEL, Christian HELFT, Dirk HOFFMANN et Gaëlle SHIFRIN.



n°32
Novembre
2015

La lettre IN2P3 Informatique

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

A LA
UNE

Ça bouge au RI3 !
Communauté



Ce n'est pas si fréquent, parlons un peu du Réseau des Informaticiens de l'IN2P3 et de l'IRFU, le RI3. Pour certains, cela reste une structure floue, sans impact clairement identifié, dont on admet volontiers ne pas bien connaître les actions, encore moins les rouages. Par contre, on lui attribue, à juste titre, l'organisation des Journées Informatique et leur succès. Cela tombe bien, les JI reviennent en 2016 (fin septembre début octobre, à Clermont) !

Tout d'abord : avouons-le, cet article est écrit à quatre mains, car Christian, qui voit se profiler une heureuse retraite à l'horizon, passe la main à Frédérique pour l'animation du RI3. Mais soyez rassurés (ou restez inquiets, c'est selon ;-), Christian reste là pour surveiller que tout se passe bien, pour encore un peu moins d'un an maintenant.

CCRI du RI3, kesako ? Loin de reposer sur une seule personne, ou deux, l'animation du réseau se discute et se met en place au sein du comité de coordination du réseau (CCRI). Depuis le printemps dernier, la petite dizaine de personnes qui compose ce comité a pris le parti de se réunir au moins une fois par mois, se voulant une structure plus réactive, à même d'interagir aussi avec le directeur du Centre de Calcul, le chargé de mission à l'informatique et le DAS (directeur adjoint scientifique) chargé du Calcul à l'IN2P3.

Pourquoi un réseau ?

Le passage de témoin pour l'animation du RI3, et la tenue du colloque de prospective « Calcul » organisé par la direction de l'IN2P3 en juin dernier se sont révélés être deux bonnes opportunités pour se poser un certain nombre de questions concernant l'identité, les finalités et l'animation du réseau. Des points de vue ont été recueillis par le biais d'une enquête reprenant les grands thèmes du colloque et évoquant la vocation du RI3. La **synthèse « Le RI3 et le colloque Calcul »** est disponible sur le site du RI3. Réseau métier ou réseau thématique : la question n'est pas tranchée. Affilié à **Resinfo**, la fédération des réseaux métiers ASR, et à **DevLog**, le réseau du Développement Logiciel Recherche et Enseignement supérieur, RI3 est plutôt considéré comme un réseau thématique ou « disciplinaire ». L'idée est aujourd'hui de conjuguer les deux points de vue : « métier » (amélioration des compétences de production de logiciel et de gestion des infrastructures informatiques) et « thématique » (prise en compte des enjeux spécifiques de nos disciplines). Il est ainsi envisagé de renforcer l'animation autour des problématiques spécifiques de l'informatique IN2P3-IRFU, par exemple en faisant du défi du parallélisme et de la vectorisation des logiciels une priorité, sans pour autant oublier la vocation de base du réseau et les attentes de la communauté qui portent essentiellement sur l'information et la formation.

Mise en place d'un groupe de travail "Services Collaboratifs à l'IN2P3"

En réponse à la **problématique exposée par Jean-René Rouet lors des JI2014**, le CCRI propose la mise en place au sein du RI3, d'un groupe de travail dédié aux Services Collaboratifs. Historiquement, les outils collaboratifs ont plutôt été déployés par le CC-IN2P3 pour couvrir ses besoins internes, certains d'entre eux ayant été ouverts après coup vers l'extérieur pour répondre aux besoins des unités ou des projets. Aujourd'hui, l'idée est d'adopter pour la mise en œuvre et la maintenance de ces outils logiciels une démarche collaborative qui ne repose pas exclusivement sur les ressources humaines du CC-IN2P3. Les premières actions attendues de ce groupe de travail sont typiquement de :

- faire le recensement des outils déjà installés au sein de l'IN2P3,
- mettre en perspective ces outils avec ceux disponibles à l'extérieur de l'IN2P3 (CNRS, Renater, Universités, CERN, etc.)
- veiller à la mise en place d'un portail documentant (fonction, accès, cible visée, etc.) les services déjà en place et réputés opérationnels en l'état,
- installer des sous-groupes de travail spécifiques à une classe de services, qui fera l'objet d'une recommandation de choix de produit, et surtout de mode de mise en œuvre et de maintenance.

Jean-René a accepté d'animer ce groupe. Il est convenu de désigner un(e) co-animateur (trice) au sein des unités autres que le CC-IN2P3 lors du lancement d'ici la fin de l'année ou le tout début de l'année prochaine.

La communication

Les deux cellules « Lettre Informatique » et « Formation & Evénements » ont également un rôle important. La première, animée par Gaëlle Shifrin et Virginie Delebarre Dutruel (toutes deux du CC-IN2P3), fait régulièrement appel à vous pour cette publication bimestrielle préparée conjointement par le Centre de Calcul et le CCRI. La seconde est là pour réfléchir aux prochains sujets de l'école IN2P3 informatique avec le chargé de formation de l'IN2P3, veiller au lancement des Journées Informatique tous les deux ans et proposer des actions d'animation « intermédiaires » entre ces Journées. Valérie Givaudan (LAL), qui a pris la suite de Françoise Virieux (APC) en septembre dernier, a réuni un groupe d'une petite dizaine de personnes en partie issues du CCRI. Parmi les nouvelles bonnes idées de ce groupe, il y a une volonté forte de valoriser le plus possible le matériel de formation dont nous disposons et de proposer des actions de formation plus courtes que les écoles et sous des formes différentes afin de toucher une plus large audience que les participants à l'école. Enfin, la maintenance et l'évolution des canaux de communication (mise à jour des listes de diffusion, maintenance et/ou évolution du site informatique.in2p3.fr, reprise des webinaires) font partie de ces « petits travaux du CCRI » toujours à l'ordre du jour ou en cours de réflexion.

Si d'aventure, vous souhaitez soutenir l'une ou l'autre de ces idées, n'hésitez pas à nous faire part de votre enthousiasme. Si ce n'est déjà fait, **rejoignez la communauté des Informaticiens de l'IN2P3 et de l'IRFU** en vous inscrivant sur le site et sur l'une des listes de discussion du RI3 (ou les deux...).

Frédérique CHOLLET (LAPP), Christian HELFT (LAL)

RI3, point de vue « métier »

<p>○ Pour l'informaticien</p> <ul style="list-style-type: none">• sortir de son « isolement »<ul style="list-style-type: none">• échanger ses connaissances, savoir-faire avec ses collègues• savoir ce qui se fait dans sa communauté<ul style="list-style-type: none">• élever le niveau de sa pratique professionnelle par d'autres moyens que la formation	<p>RI3</p>	<p>○ Pour la tutelle</p> <ul style="list-style-type: none">• avoir une vision du potentiel de la communauté• pouvoir faire appel à un « vivier » de compétences• disposer de « remonteurs d'alertes technologiques »• construire des plans de formation au plus près des besoins
---	------------	---



RI3, point de vue « thématique »

<p>○ Pour l'informaticien</p> <ul style="list-style-type: none">• interagir au-delà de son cercle habituel• comprendre les finalités de l'informatique IN2P3, les intégrer dans la conduite de ses activités• discerner les enjeux, les priorités de la discipline face au foisonnement technologique• savoir ce qui se passe à l'international	<p>RI3</p>	<p>○ Pour la tutelle</p> <ul style="list-style-type: none">• avoir une vision du potentiel de la communauté• pouvoir faire appel à un « vivier » de compétences spécifiques• disposer d'informations utiles pour la conduite de l'informatique IN2P3• coordonner et soutenir des propositions et initiatives d'envergure nationale ou internationale
--	------------	---



n°32
Novembre
2015

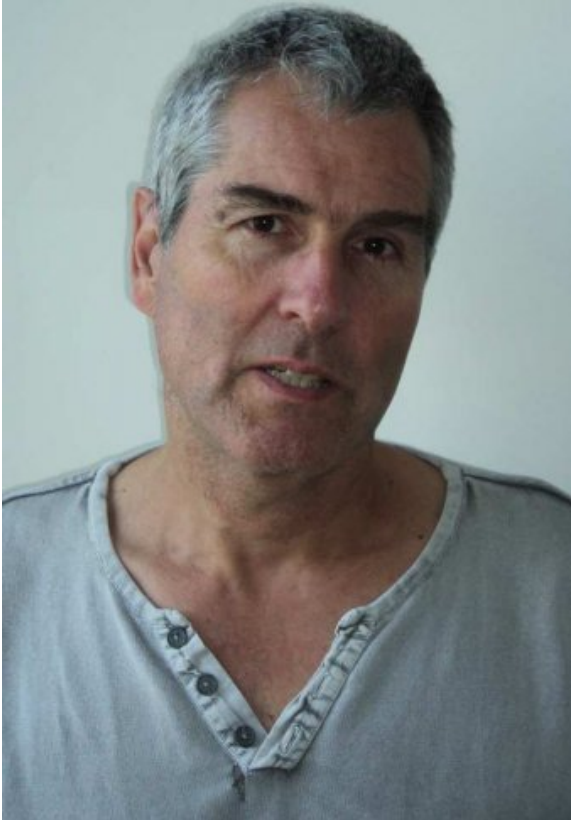
La lettre IN2P3 Informatique

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



"il faut se garder de penser que nous sommes à part en matière d'informatique"

Thierry Ollivier, chargé de mission formation à l'IN2P3



- Tu as exercé différentes fonctions à l'IN2P3 et participé, en 2004, à la création du Réseau des Informaticiens IN2P3 et IRFU. Peux-tu nous expliquer ce qui a conduit à la genèse du RI3 ?

La genèse de ce réseau a été une longue évolution. La structuration en réseau ou groupe de travail va de soi lorsqu'une communauté technique est bien ciblée et tournée vers un objectif technique commun bien identifié. Historiquement (sans vouloir jouer les anciens), j'ai connu le groupe de travail « VAXpopuli », très actif à l'époque où tous les labos avaient une machine centrale VAX/VMS. Des groupes s'étaient aussi créés autour de certains protocoles réseau. Plus récemment, le réseau technique LCGFR est un parfait exemple de collaboration inter-labos lié à un objectif dédié. De même que le groupe « W2K » autour d'une forêt Active Directory IN2P3. De telles communautés d'informaticiens peuvent avoir un impact important dans l'évolution technique de nos infrastructures. On peut penser par exemple que le petit groupe qui s'était créé autour de la vidéoconférence H323 a contribué de façon importante à sa généralisation dans les labos et à l'achat du premier MCU par le CC-IN2P3. On peut aussi citer la collaboration autour des serveurs de fichiers NAS NetApp qui permettait d'obtenir un modèle de coût unique à l'IN2P3.

Des formes de réseau de métier ont donc toujours existé, dans un domaine pour lequel les évolutions techniques permanentes rendent indispensables les échanges entre informaticiens. Ceux-ci sont probablement plus faciles chez les ASR que chez les développeurs, souvent tenus par les contraintes de l'expérience elle-même.

La diversité des techniques (online, offline, langages...) rend également plus difficile d'identifier des collègues confrontés aux mêmes problèmes. Et pourtant dans ce domaine il y a vraiment

En plus, le comité d'organisation issu du CCRI essaie de les renouveler en permanence avec par exemple les présentations express (5 min) qui permettent de prendre ensuite des contacts pour aller plus loin. Les webinaires ont été aussi une vraie initiative nouvelle, même si ensuite les faire vivre année après année reste un challenge. Les domaines techniques et les groupes de travail potentiels ont été bien identifiés, voire structurés. Et notre réseau national est représenté dans RESINFO (réseau des réseaux, qui fédère plutôt des réseaux régionaux) et à travers lui, auprès de la Mission pour l'interdisciplinarité.

Tout n'est pas parfait bien sûr, par exemple les listes ASR et DEV créées pour notre communauté ne sont que très peu actives. Notre réseau est forcément concurrencé par les réseaux régionaux qui, eux aussi, sont actifs et organisent des actions. Mais je reste persuadé que notre réseau national RI3 reste pertinent et cohérent avec nos disciplines scientifiques et leurs exigences en informatique.

- Tu as également été chef du service informatique à l'IPNL durant 23 ans. Quel regard portes-tu sur l'informatique à l'IN2P3 ?

Objectivement, je pense qu'il faut se garder de penser que nous sommes à part en matière d'informatique. J'ai le sentiment que d'autres instituts ont beaucoup progressé ces dernières années, par exemple l'informatique pour la biologie ou pour les SHS. Le programme du colloque de restitution du défi Mastodons en décembre montre par exemple que la majorité des disciplines sont engagées dans des projets de gestion des données autour du « big data ». Nous n'avons pas le monopole des grands projets.

Mais à l'IN2P3, nous avons à mon sens trois avantages concernant notre métier : le fait que nos thématiques scientifiques nous imposent d'être présents sur la totalité des domaines techniques de l'informatique, intégrant souvent un aspect de R&D ; une structuration des services et de la communauté (on revient au point précédent) qui fait que nos jeunes informaticiens sont accompagnés dans leurs projets. Cette structuration des équipes tire notre discipline vers le haut ; Enfin, la présence d'un centre de calcul national qui permet d'appuyer des initiatives des labos en pouvant les consolider en outils nationaux.

En tant que chef de service, j'ai toujours insisté ou été favorable à ce que mes collègues participent aux réseaux, événements ou projets collectifs et n'hésitent pas à contacter nos autres labos. A l'IN2P3, un informaticien qui le souhaite n'est pas isolé dans son travail.

Ces éléments nous rendent, je pense encore attractifs en BAP E, et par exemple dans mon service je n'ai jamais connu de difficultés pour nos recrutements en NOEMI ou C.E..

Et pour avoir, de par ma mission formation, été impliqué dans l'organisation des trois dernières écoles informatiques (ANF) (qui s'appuient pour la partie TP sur nos infrastructures et en particulier le CC-IN2P3), je peux vous dire que les participants d'autres instituts que nous accueillons me font souvent part de leur impression d'une informatique structurée et attrayante.

- Concernant la formation permanente, vois-tu des spécificités propres à l'IN2P3 concernant l'informatique et quels sont les axes de réflexion ?

Déjà le fait que le CCRI ait une cellule « Formation et événements » fait qu'en tant que chargé de mission j'ai comme interlocuteur une structure représentative de la communauté, qui a vocation à être le comité de pilotage des écoles. Nous

matière à mutualiser des briques de base et du savoir-faire. Des évènements tels que nos JI tous les 2 ans, ou les JDEV, montrent que tous nos métiers peuvent bénéficier de rencontres régulières. Il faut donc rendre hommage à Christian Helft pour sa persévérance à avoir souhaité structurer et animer la totalité de notre communauté des informaticiens. Au-delà des groupes et évènements existants, il était important d'asseoir la reconnaissance par l'IN2P3 de nos métiers au travers d'une organisation plus formalisée qui puisse aider à fédérer des groupes de travail, initier des domaines de collaboration, lancer des prospectives, gérer des évènements et de la communication.

Aux JI 2004, avec Christian et Frédérique (Chollet), nous en avons beaucoup discuté, et tracé dans un document les contours de ce que pourrait amener une plus grande coordination au niveau national. Ensuite, nous avons aidé Christian à soutenir le projet auprès de l'IN2P3. Jusqu'au moment où en 2007, avec Cristinel Diaconu comme chargé de mission informatique, le CCRI s'est mis en place. Le RI3 avait de fait une reconnaissance. Personnellement, deux points me tenaient à cœur, justement cette lettre d'information régulière et également que le CCRI crée un lieu de discussion entre l'ensemble des labos et le CC-IN2P3.

Le fait d'avoir accepté d'être responsable technique de l'IPNL de 2007 à 2013 puis chargé de mission formation de l'IN2P3 depuis 2013 fait que j'ai suivi d'un peu plus loin l'activité du RI3. Je retrouve depuis 2013 le CCRI à travers sa cellule « Formation et évènements » pour ma mission de formation permanente qui inclut la définition des écoles informatiques.

- Comment as-tu vu évoluer le réseau depuis ses débuts ?

Il me semble objectivement que le CCRI a plutôt bien évolué depuis 2007. Déjà il est resté « vivant » et il est l'interlocuteur reconnu dans plusieurs domaines. La lettre informatique continue à un rythme de 3 ou 4 par an ; les JIs me semblent vraiment pérennisées, même si les problèmes budgétaires restent un souci potentiel ; et pour ma part, je vois que la cellule « Formation et évènements » bouillonne d'idées. Concernant les JI (je crois avoir participé à toutes depuis 1999), je suis toujours étonné de l'affluence et également de voir comment, dans tous les domaines techniques, les participants sont intéressés de découvrir ce que les autres ont fait.

[1] MOOC : Massive Open Online Course

sommes soutenus par l'IN2P3 pour commanditer une école informatique par an et je fais remonter à la direction technique d'autres sujets que j'essaie d'identifier par des discussions avec la communauté.

Ainsi, pour 2016, nous demandons au CNRS le budget pour piloter trois écoles, l'école annuelle qui serait autour de la programmation hybride utilisant tous les modes possibles d'accélération matérielle offerts par les machines actuelles, une école GEANT4, et une école autour du passage de nos labos à IPV6. Auxquelles il faut rajouter une ANF sur ATRIUM qui n'est pas strictement une école informatique, mais est néanmoins un développement informatique interne à notre Institut et concerne pas mal de personnes de nos services.

Il existe des réseaux de métier qui proposent des écoles à travers la mission interdisciplinaire (MATHRICE, groupe calcul), mais cette structuration de la formation informatique à un niveau « Institut » est je pense unique. La cellule du CCRI a plein d'idées nouvelles : en vrac, déjà essayer de rejouer l'école informatique annuelle sous une forme en non-présentiel, à mi-chemin entre une école et un MOOC [1]. C'est un vrai challenge technique de faire cette transformation mais cela permettrait qu'un tel investissement technique ne soit pas « one shot ».

On envisage également de créer en 2016 de zéro une école en mode « à distance » (par exemple 8 demi-journées avec cours en mode visio et outils de suivi des TP à distance). Enfin, on doit penser aussi à des sessions de formation plus courtes (une journée) sur un thème très ciblé. Sans parler bien sûr de voir comment la formation peut être acteur dans le suivi de MOOC par nos collègues ; nous avons des discussions au niveau du CNRS sur ce point, mais ils ne sont toujours pas reconnus pour l'instant comme action de formation. On peut néanmoins en interne échanger sur ceux-ci, essayer de cibler les plus pertinents, se conseiller entre labos...

Sur ces nouveaux modes de formation, il faut définir ce que peut amener la cellule du CCRI et ce que peut amener la formation permanente de l'IN2P3. Nous travaillons ensemble à le définir.

Propos recueillis par Gaëlle SHIFRIN



n°32
Novembre
2015

La lettre IN2P3 Informatique

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



CentOS Seven, changement de cap



CentOS

Le système d'exploitation Scientific Linux (SL) [1] est utilisé depuis 2003 par la communauté HEP [2] et par l'IN2P3 pour ses environnements de calcul. Scientific Linux est l'un des clones de la très populaire distribution commerciale Red Hat Enterprise Linux (RHEL) [3]. SL est né d'une initiative du laboratoire américain Fermilab [4], initiative à laquelle le CERN [5] a apporté une forte contribution.

Depuis la version 3 de SL, très instable à ses débuts, à la version 7 en 2015, un long chemin a été parcouru. La qualité du conditionnement du système d'exploitation n'a cessé de s'améliorer pour arriver aujourd'hui à un niveau comparable à bon nombre de distributions commerciales.

Toutefois, des interrogations émergent sur le cap à tenir. Si le conditionnement de Scientific Linux n'a plus à rougir face à ses compétiteurs, la politique de fourniture des mises à jour semble encore perfectible. Certains la jugent en retrait par rapport à celle de la distribution CentOS [6], autre clone populaire de RHEL. Des interrogations sur la pérennité de l'offre Scientific Linux sont également très présentes.

Le rapprochement entre la société Red Hat [7] et le projet CentOS en janvier 2014, a eu pour effet d'amplifier les interrogations sur l'avenir de SL et sur son possible abandon dans la communauté HEP.

Certains sites WLCG [8] avaient déjà fait le choix de CentOS au détriment de Scientific Linux dès le début. Le CERN quant à lui, après l'avoir annoncé en 2014, acte définitivement l'abandon de SL avec la bascule à CentOS 7 en avril 2015 [9].

Utilisateur de Scientific Linux de la première heure et de sa version 6 aujourd'hui, le Centre de Calcul de l'IN2P3 (CC-IN2P3) [10], n'a pas échappé à ces interrogations « métaphysiques ». Après un temps de réflexion, il a finalement été décidé que la version 7 de son système d'exploitation serait CentOS plutôt que Scientific Linux. Si CentOS 7 est déployé sur les nouveaux serveurs depuis le printemps 2015, l'introduction de ce nouveau système d'exploitation dans la ferme de calcul et sur la plateforme d'accueil interactif du CC-IN2P3 sera programmée en 2016.

Dans la mesure où CentOS et SL sont des reconditionnements de RHEL, ce changement de distribution ne devrait avoir que très peu de répercussions sur l'environnement proposé par le Centre de Calcul à ses utilisateurs. Le Centre de Calcul de l'IN2P3 tient à remercier le laboratoire Fermilab et tous ses collaborateurs impliqués dans le projet Scientific Linux.

Ces 13 années de cheminement commun ont permis au CC-IN2P3 de tenir sa position de Tier1 WLCG et de préparer le futur. Nous ne pouvons qu'être reconnaissants et souhaiter bonne route au projet Scientific Linux.

Dominique CLEMENT (CC-IN2P3)

[1] <http://www.scientificlinux.org>

[2] HEP : High Energy Physics / Physique des Hautes Énergies

[3] <http://www.redhat.com>

[4] <http://www.fnal.gov>

[5] <http://www.cern.ch>

[6] <http://www.centos.org>

[7] <https://www.redhat.com/en/about/pre...>

[8] WLCG : Worldwide LHC Computing Grid <http://wlcg.web.cern.ch/>

[9] <http://linux.web.cern.ch/linux/centos7/>

[10] <http://cc.in2p3.fr>



n°32
Novembre
2015

La lettre IN2P3 Informatique

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



Dans les coulisses de l'école informatique



Chaque année, l'IN2P3 organise une école informatique (voir l'interview de Thierry Ollivier dans ce numéro). La préparation de la dernière en date [1] s'est révélée une expérience très enrichissante pour tous ses acteurs. Le thème était très ambitieux, mais l'équipe rassemblée à cette occasion, emmenée par le tandem Thomas Lallart de l'INRA et Foudil Brétel du CC-IN2P3, a relevé le défi, avec un certain panache, je tiens à le souligner.

Bien sûr, l'engagement, les compétences et le talent des intervenants ont été les principaux ingrédients de la réussite de cette petite aventure humaine. Mais je voudrais citer trois outils qui ont rendu possible celle-ci, vécue par une équipe dispersée géographiquement.

GitLab, le gestionnaire de dépôt de code partagé par un groupe de développeurs, a permis à chaque membre de l'équipe de connaître dès la conception, commenter, modifier ou compléter la production de ses collègues. Pas seulement du code, mais aussi et en l'occurrence surtout, des énoncés des exercices et des présentations.

Rappelons que le CC-IN2P3 offre un service GitLab national, et que plusieurs labos de l'IN2P3 ont mis en place une instance locale.

Rendez-vous, le système de vidéoconférence « personnelle et instantanée » (ma définition du service, pour faire court), offert par Renater, a remplacé une multitude de mails et de coups de téléphone, de façon beaucoup plus efficace. La légèreté de mise en œuvre de l'outil a permis de l'utiliser plusieurs fois par semaine, souvent de façon improvisée, et sa capacité de partage d'écran très efficace s'est révélée précieuse.

Un système de « **salon de discussion textuelle** » (chat en français...), dans lequel chacun entrait automatiquement dès la mise en route de son poste de travail, a permis encore un autre mode d'échanges, comme si toute l'équipe était réunie dans la même pièce, même pendant les périodes de travail en dehors de la préparation de l'école. Une fois le mode d'emploi d'un tel outil bien compris, celui-ci permet d'accéder aux avantages d'une interaction permanente sans être trop intrusif.

Chaque organisateur potentiel d'une école, dont les intervenants doivent interagir fortement lors de la préparation, pourrait probablement tirer profit de ces outils désormais largement accessibles (ou de leurs équivalents). Associés à une mailing-list des intervenants et à un site Indico pour publier le programme et les présentations y compris avec les modifications de dernière minutes, ils constituent un socle solide de construction d'une école ou de tout événement similaire.

Un autre aspect digne d'être mentionné est l'impact réciproque de cette école et des infrastructures informatiques de l'IN2P3/IRFU. Si à l'avenir vous voyez apparaître dans votre paysage informatique des services Sonar ou autres repositories

GitLab, la qualité de code et la dette technique avec Sonar, l'intégration continue avec Jenkins, l'analyse de logs et le packaging de code et le déploiement avec Docker. Pendant toute la durée de l'école, un code en Java permettant de lancer une application web d'analyse de données a été utilisé afin d'illustrer l'intérêt des différents outils présentés.

Le lundi, Thomas Lallart (DSI de l'INRA) et coordinateur de la partie technique de l'école, nous a présenté le **cycle de vie du logiciel** et les outils qui y sont associés. Nous avons vu aussi des notions importantes comme la **qualité logicielle** et la **dette technique**. Quelques cas concrets de ces bonnes pratiques ont conclu la journée.

La journée du mardi était consacrée aux **tests logiciels** avec en introduction une présentation théorique d'une chercheuse en informatique, Marie-Claude Gaudel (Université Paris Sud). L'outil utilisé ensuite en travaux pratiques était **Sonar**, un logiciel de gestion de qualité de code très utile (utilisation « stand-alone » ou depuis un serveur). Sonar permet de vérifier si les conventions de codage sont bien respectées et il calcule la dette technique en se basant sur les exigences des développeurs ce qui en fait un outil indispensable au développement.

Nous avons utilisé **Git** le mercredi avec Benoît Bayol (ECP) afin de pouvoir prendre en main ce système de gestion de version très souple mais aussi assez complexe. À travers l'interface **GitLab** et l'API en ligne de commande, nous avons créé des projets, géré différentes branches et fusionné celles-ci afin de voir les possibilités multiples de Git. Antoine Pérus (LAL) nous a aussi présenté le concept de flux de travail (« **Workflow** ») pour le développement et particulièrement la gestion des branches à l'aide d'outils spécifiques. Avec Foudil Brétel (CC-IN2P3), nous avons pratiqué Jenkins, un outil d'intégration continue qui exécute automatiquement dans un projet la construction du code et effectue les tests unitaires, vérifie la qualité de code ou encore crée la documentation. Nous avons configuré **Jenkins** (sur un serveur) pour qu'il construise le code Java compilé avec Maven et nous avons pu ainsi découvrir la variété de « **plug-ins** » et donc de possibilités offertes par cet outil.

Le jeudi était consacré à l'**analyse de logs** et à **Docker**. Fabien Wernli (CC-IN2P3) nous a présenté la gestion des logs au CC-IN2P3. Nous avons utilisé ensuite une librairie Java de log générique qui pilote un logger. Le déploiement avec Docker a été présenté par Sébastien Binet (LPC Clermont). Docker permet d'instancier des conteneurs (machines virtuelles légères) à la demande en local (dans le TP) et sur les infrastructures de Cloud (OpenStack, AWS...). Sur Mac OS X et Windows, l'utilisation de Docker se fait à l'heure actuelle via une machine virtuelle dans laquelle les conteneurs vont être créés. Nous avons utilisé les fonctions de base de Docker comme la récupération d'une image Docker depuis le DockerHub (dépôt public), son utilisation à travers un conteneur et la création d'image via les Dockerfiles. Mais nous avons pratiqué des fonctions plus avancées en compilant le code Java dans des conteneurs et en manipulant le concept de « **registry** » qui permet de créer un dépôt privé en local ou distant.

Le vendredi concluait cette semaine très riche en technique avec une présentation de Fabien Wernli et Jean-René Rouet (CC-IN2P3) sur l'importance des messages d'erreurs dans les logs et l'utilisation d'**ElasticSearch** au CC-IN2P3.

Cette école était une grande réussite car, d'une part, les participants sont repartis avec des outils prêts à être utilisés pour l'amélioration de leurs codes et un grand nombre de bonnes pratiques à mettre en place dans leurs projets et, d'autre part, les intervenants très chaleureux et disponibles ont vraiment pu transmettre leurs connaissances dans une ambiance joyeuse et créative. Un grand merci à toute l'équipe.

Docker, je me plais à penser que l'école n'y aura pas été pour rien.

Je voudrais conclure en revenant au point de départ : sans l'implication profonde, je préférerais même dire sincère, des intervenants, point d'école de qualité. Et, au-delà de mon coup de chapeau pour toute l'équipe, s'il faut n'en citer qu'un, je remercie particulièrement Thomas pour avoir mis au service de cette édition toutes ses compétences, son expérience assez unique de mise en place d'un véritable suivi de qualité en informatique, son engagement personnel et son remarquable et remarqué relationnel.

Pour l'autre côté du rideau, je laisse la parole à Cécile, à qui nous avons demandé un petit compte rendu de l'école, puis Nathalie et Raphaël pour leur vécu, comme participants en dehors de la communauté IN2P3-IRFU.

Cécile Cavet (APC) :

L'école informatique de l'IN2P3 a eu pour thème cette année « **Les outils de mise en production du logiciel** » et elle s'est déroulée à Lyon du 28 septembre au 2 octobre 2015. Elle s'adressait à tout développeur ayant déjà parcouru le cycle du développement logiciel et qui voudrait faire évoluer celui-ci en se basant sur de meilleures pratiques. Pendant 5 jours, les participants hébergés au Centre Jean Bosco sur la colline de Fourvière ont suivi des présentations et des travaux pratiques de grande qualité au CC-IN2P3.

Les sujets abordés étaient le cycle de vie du logiciel et ses outils : le système de gestion de version Git avec l'interface

[1] <https://indico.in2p3.fr/event/11728/>

Nathalie Poulet (LMD) :

Je suis Assistant Ingénieur à l'Institut Pierre Simon Laplace (INSU). Je devais débiter le développement d'un gestionnaire de flux de données lorsque j'ai vu l'annonce de l'école "les outils de mise en production du logiciel". Dans notre équipe, nous avons la volonté de mettre en oeuvre certaines bonnes pratiques de développement dès le début de ce projet ; le programme de l'école répondait à cet objectif. Durant cette formation, j'ai apprécié la diversité des intervenants (recherche, secteur privé, ingénieurs) ainsi que l'alternance cours et TP pour se confronter aux outils présentés. L'appartenance de la majorité des intervenants à des laboratoires du CNRS et de l'INRA est un vrai atout car la formation était orientée sur des problématiques communes à nos pratiques professionnelles. J'ai également beaucoup aimé la richesse des échanges avec les participants sur nos expériences réciproques.

Raphaël Tournoy (CCSD) :

Cette école m'a permis d'avoir un bon aperçu des pratiques et outils liés au cycle de vie du logiciel. Les travaux pratiques m'ont permis de prendre le temps d'expérimenter de nouveaux outils que je n'avais pas pris le temps de tester. J'espère que la plupart des méthodes abordées vont progressivement être mises en oeuvre dans les développements de notre unité. Merci aux organisateurs, aux intervenants et aux participants pour cette formation dans une ambiance conviviale.

Christian HELFT (LAL)



Docker & HENP



Le développement logiciel en physique des hautes énergies – en particulier pour les expériences LHC – met en jeu l'assemblage et l'intégration d'un grand nombre de bibliothèques (externes ou internes, scientifiques ou générales), qui doivent être ensuite compilées, testées et installées sur les machines de production.

De plus, et à cause des ressources limitées disponibles, les logiciels des expériences ne sont testés et déployés que sur un jeu très limité de plateformes, architectures, chaînes de compilation et systèmes d'exploitation. Mécaniquement, l'installation d'un environnement de travail sur une machine de développement avec un système légèrement différent (Debian/CentOS/...), ou bien sur une machine de production exploitant une nouvelle architecture, est généralement un processus itératif chronophage, surtout lorsque les performances natives (sans machine virtuelle) sont requises.

En effet, les machines virtuelles (VM) ont grandement facilité le développement et le déploiement d'applications d'une part, ainsi que la gestion et l'optimisation de l'utilisation de clusters de calcul d'autre part. Cependant, la virtualisation de l'environnement de travail s'est effectuée au prix de performances diminuées par rapport aux performances natives.

Si les VMs offrent une virtualisation au niveau d'une machine entière, les conteneurs offrent quant à eux une virtualisation de l'environnement au niveau du système d'exploitation. Les conteneurs semblent donc mieux répondre aux besoins de la physique nucléaire et des hautes énergies (HENP), l'environnement de production étant essentiellement basé sur des clusters Linux. LXC [1] et OpenVZ [2] ont été les premiers à introduire les conteneurs dans l'écosystème Linux, mais Docker [3] est le projet qui les a réellement popularisés et démocratisés.

Cet article explore les possibles applications des conteneurs Docker aux environnements de travail typiques dans HENP.

Technologies utilisées

Docker est un projet *open source* écrit en Go [4] pour empaqueter, distribuer et lancer n'importe quelle application à l'intérieur d'un conteneur.

Docker utilise les espaces de nommage (*namespaces*) Linux, les *cgroups* [5] et un système de fichiers permettant la fusion de points de montage (*unioning*) pour isoler les processus. Les images de conteneurs sont similaires aux images de machines virtuelles, mais partagent le noyau Linux avec la machine hôte. Cette configuration met à disposition un système plus léger et permet de provisionner des images de conteneurs en quelques secondes – à comparer aux plusieurs minutes pour les VMs – mais permet également de lancer plusieurs centaines de conteneurs sur une machine de bureau typique.

Grâce aux *cgroups*, les conteneurs peuvent avoir leurs propres interfaces réseau et sont familièrement vus comme un *superchroot* [6], sans avoir recours à une quelconque émulation de matériel – permettant ainsi d'atteindre les performances optimales de la machine physique. Docker utilise également un système de fichiers à plusieurs couches via un système de

Création d'images personnalisées

Un utilisateur peut créer de nouvelles images de manière interactive. Pour cela, il lui suffit de :

- lancer un conteneur à partir d'une image de base,
- lancer des commandes de manière interactive à l'intérieur du conteneur,
- et sauvegarder l'état actuel du conteneur dans une nouvelle image, comme montré dans la figure 6.

```
>>> docker run -i -t ubuntu bash
root@5ad62f3a9b4b:/# apt-get install -y memcached
[...]
Unpacking memcached (1.4.14-0ubuntu9) ...
root@5ad62f3a9b4b:/# exit
exit

>>> docker ps -l
CONTAINER ID        IMAGE                COMMAND              CREATED
5ad62f3a9b4b      ubuntu:latest       "bash"               3 minutes ago

>>> docker commit 5ad62f3a9b4b binet/memcached
220c5747349a7ec1e16297eecdffileacc277b9a9a6a149e386aff9f63bec868e

>>> docker images
REPOSITORY          TAG                IMAGE ID            CREATED             VIRTUAL
SIZE
binet/memcached     latest            220c5747349a       4 minutes ago     190 MB
ubuntu              latest            07f8e8c5e660       5 days ago        189.3 MB
centos              latest            f044297e2dd0       2 weeks ago       215.7 MB
fedora/apache       latest            963668e7ae33       2 weeks ago       627.1 MB
ubuntu              12.10            c5881f11de09       10 months ago     172.1 MB
```

Figure 6 `docker run` lance la commande `bash` dans le conteneur, en mode interactif (-i) avec un pseudo-terminal TTY (-t). Une fois que le conteneur est dans l'état voulu (les paquets nécessaires installés, les applications configurées correctement, etc.), il peut être sauvegardé comme une nouvelle image, appelée `binet/memcached` dans cet exemple.

La création d'images en interactif est très utile pendant le développement ou le débogage du processus de création. Cependant, une interface de scriptage est nécessaire pour la reproductibilité du processus de création ainsi que pour son passage à l'échelle. Le projet Docker a donc introduit le concept du fichier `Dockerfile` et défini ses spécificités. Les fichiers `Dockerfile` sont ainsi l'équivalent des `Makefile` [11] pour la création de nouvelles images. La syntaxe de ces fichiers ressemble à celle des scripts shell, avec quelques mots clés listés et décrits par `Dockerfile syntax reference` [12].

Le fichier `Dockerfile` équivalent aux commandes de la figure 6 est reporté dans la figure 7. La nouvelle image correspondant à la recette décrite par le fichier `Dockerfile` est créée en lançant la commande `docker build` depuis le répertoire contenant le fichier `Dockerfile`.

```
>>> cat Dockerfile
## create a memcached image
FROM ubuntu
MAINTAINER me@example.com

RUN apt-get install -y memcached

>>> docker build --tag=binet/memcached .
Sending build context to Docker daemon 2.048 KB
Sending build context to Docker daemon
Step 0 : FROM ubuntu
----> 07f8e8c5e660
Step 1 : RUN apt-get install -y memcached
----> Running in 7f8349773ece
[...]
----> aac9f626b9fb
Removing intermediate container 7f8349773ece
Successfully built aac9f626b9fb

>>> docker images
REPOSITORY          TAG                IMAGE ID            CREATED             VIRTUAL SIZE
binet/memcached     latest            aac9f626b9fb       About a minute ago 190 MB
```

Figure 7 Un exemple simple de fichier `Dockerfile`, scriptant la création d'une nouvelle image à partir de l'image `ubuntu`, dans laquelle l'application `memcached` est installée.

Cas d'usages

Docker est très utile pour un certain nombre de *workflows* typiques dans HENP. Docker permet en effet :

- l'encapsulation de la construction de l'entièreté de la pile logicielle d'une expérience, en s'assurant qu'il n'y a aucune dépendance externe implicite cachée,
- l'installation et le déploiement d'une pile logicielle déjà compilée sur un nombre arbitraire de nœuds et de sites, ainsi qu'une rapide mise en production de cette pile logicielle,

fichiers d'unioning (UnionFS [7], AUFS [8]) ou via *device mapper* [9] pour les noyaux Linux sans cette fonctionnalité. Chaque couche du système de fichiers est montée au-dessus des précédentes. La première couche est appelée image de base et contient la collection initiale de fichiers et répertoires mise à disposition par une distribution Linux donnée (Ubuntu, RHEL, Fedora, etc.), distribution qui n'est pas nécessairement la même que celle de la machine hôte. Chaque couche est en lecture seule (seule la dernière couche est modifiable) et seules les différences avec la couche précédente sont stockées sur disque. Les différentes couches sont indexées via un hash à la git (somme de contrôle) et peuvent être partagées et réutilisées par différentes images : cette stratégie permet d'optimiser les ressources de stockage et les performances de provisionnement de disques.

- la distribution simple et rapide d'environnements de développement.

Docker permet également une répartition plus souple des responsabilités entre les équipes de développement et de production, en adéquation avec la mouvance *DevOps* [13].

L'équipe de développement peut ainsi choisir la distribution Linux adéquate, le portefeuille de dépendances externes (ainsi que leur version), son cadriciel, etc. sans impact pour l'équipe de production (autre que l'adoption des conteneurs comme outil de mise en production). De son côté, l'équipe de production peut se focaliser sur la gestion des machines, la gestion des logs, la politique de sauvegarde des données, le *monitoring*, etc. sans interférer avec le processus de développement.

Dans cette optique, un certain nombre d'images de base ont été créées spécifiquement pour la communauté HENP et ont été collectées dans le dépôt [gi-thub.com/hepsw/docks](https://github.com/hepsw/docks). Entre autres :

- Scientific Linux CERN 5 ([hepsw/slc5-base](https://github.com/hepsw/slc5-base)),
- Scientific Linux CERN 6 ([hepsw/slc6-base](https://github.com/hepsw/slc6-base)) et,
- CERN CentOS 7 ([hepsw/cc7-base](https://github.com/hepsw/cc7-base)).

Dans ce dépôt se trouvent également les fichiers Dockerfile nécessaires pour la création d'images contenant l'installation binaire du cadriciel Gaudi [14] ([hepsw/lhcb-gaudi](https://github.com/hepsw/lhcb-gaudi)) et la suite logique d'analyse de l'expérience LHCb, Da Vinci ([hepsw/lhcb-davinci](https://github.com/hepsw/lhcb-davinci)). Ces images n'ont cependant pas été publiées sur le Docker Hub à cause de leur taille (O(10GB)). Il serait sans doute intéressant d'avoir un *hub* dédié à la communauté HENP, gérant de façon transparente les certificats grille.

Ce dépôt sert également des images contenant le démon CVMFs correctement installé, configuré et prêt à distribuer la pile logicielle de différentes expériences (ATLAS, CMS, LHCb et LSST).

Conclusions et perspectives

Cet article a présenté Docker et quelques-unes des applications possibles dans la communauté HENP. La "*conteneurisation*" de piles logicielles HENP est faisable, et devrait se généraliser et s'accélérer dans le futur. Les conteneurs ne présentent pas de performances dégradées par rapport aux performances natives, et ce, même pour des applications gourmandes en CPU et E/S [15] [16]. De plus, Docker (ou toute autre technologie de virtualisation légère) peut potentiellement faciliter l'optimisation de l'utilisation des clusters de calcul Linux, sans avoir recours aux machines virtuelles.

Un autre usage intéressant de conteneurs est l'emballage d'application (et ce sans couplage avec le gestionnaire de paquets de la distribution Linux sous-jacente) ainsi que la préservation des données et de la pile logicielle d'une expérience, permettant la réanalyse ou la réinterprétation des résultats plusieurs années après la fin de cette expérience.

Le paysage de la virtualisation légère est encore en pleine évolution et maturation. Des initiatives telles qu'appc [17], une standardisation du format de conteneurs pour une meilleure interopérabilité, lancée par un concurrent de Docker, rkt [18], ou bien oci [19], l'Open Container Initiative, sont à surveiller. En effet, rkt semble en bonne voie de concurrencer fortement Docker sur le front de la sécurité, ayant un code moins monolithique (et plus facilement auditable, avec des privilèges plus restreints) que ce dernier.

Une saine compétition semble s'engager entre les différents acteurs de la virtualisation légère et devrait profiter à la communauté.

Sébastien BINET (LPC Clermont)

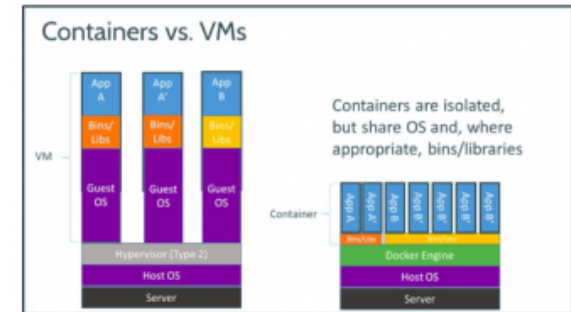


Figure 1 Comparaison entre machines virtuelles (à gauche) et conteneurs (à droite). La virtualisation légère à base de conteneurs repose sur le système d'exploitation pour l'isolation plutôt que de virtualiser tout le matériel.

Images Docker

Les images Docker sont créées à partir d'une image de base. Docker met à disposition un nombre important d'images de base officielles (ubuntu, centos, etc.). Ces images peuvent être téléchargées via la commande **docker pull** comme illustré dans la figure 2.

```
>> docker pull ubuntu
latest: Pulling from ubuntu
e9e06b06e14c: Pull complete
a82efaa99f9: Pull complete
37bea4ee0e81: Pull complete
07f8e85e660: Already exists
Digest: sha256:8126991394342c2775a9ba4a843869112da8156037451fc424454db43c25d8b0
Status: Downloaded newer image for ubuntu:latest
```

Figure 2 `pull` télécharge une image docker depuis le registre (aussi appelé docker hub).

La liste des images locales peut être consultée en lançant la sous-commande **docker images**, comme montré dans la figure 3.

```
>> docker images
REPOSITORY          TAG         IMAGE ID      CREATED       VIRTUAL SIZE
ubuntu              latest     07f8e8c5e660 3 days ago   188.3 MB
centos              latest     fd44297e2ddb 2 weeks ago  215.7 MB
ubuntu              12.10     c5881f11ded9 10 months ago 172.1 MB
```

Figure 3 `images` affiche la liste des images locales, téléchargées depuis le hub ou créées localement.

Il est également possible de lancer des conteneurs en tâche de fond, le cas typique étant les (micro)services, *daemons* et serveurs web. La syntaxe est donnée dans la figure 4.

```
>> docker run -d ubuntu sh -c 'while true; do echo "hello"; sleep 1; done;'
0ac942723c259a4963e0e2ff04d57e9bf8ad28e72158a23111d8f3718d960e

>> docker ps
CONTAINER ID   IMAGE      COMMAND                  CREATED       STATUS
0ac942723c25   ubuntu:latest  "sh -c 'while true  6 seconds ago   Up 3 seconds

>> docker attach 0ac942723c25
hello
hello
hello
hello
...
```

Figure 4 `docker run` en tâche de fond. Ici, la commande `sh` est lancée dans un conteneur basé sur l'image ubuntu.

Comme la commande est lancée en tâche de fond, il faut attacher les E/S du terminal au conteneur (0ac942723c25) pour voir les messages s'afficher. La gestion d'un conteneur peut également être effectuée via les sous-commandes **start/stop/restart**.

Les images Docker peuvent être consultées, publiées et téléchargées depuis le Docker Hub [10] : un registre global d'images officielles et d'images créées par la communauté. Cet index d'images peut être consulté depuis la ligne de commande, comme dans la figure 5, ou bien depuis le web : <https://hub.docker.com>.

```

>> docker search apache
NAME          STARS     OFFICIAL   AUTOMATED
coscat        131      [OK]
tutum/apache-php  71      [OK]
httpd         50      [OK]
maven         32      [OK]
fedora/apache  30      [OK]
[...]

>> docker pull fedora/apache
Pulling repository fedora/apache
96366e7af33: Download complete
96366e7af33: Pulling image (latest) from fedora/apache
3d26c4813f: Download complete
Status: Downloaded newer image for fedora/apache:latest

>> docker run -d -p 80 fedora/apache
128d9712c922aab640a68d561ec35f5c17a889e136be7983804035333264d92

>> docker ps
CONTAINER ID   IMAGE          COMMAND                  PORTS
128d9712c922   fedora/apache:latest  "/run-apache.sh"        0.0.0.0:32768->80/tcp

>> curl localhost:32768
Apache

```

Figure 5 `docker search` consulte le registre, cherchant des images dont le nom ou la description contient la chaîne de caractères demandée. L'image `fedora/apache` expose un serveur web Apache sur le port 80 qui doit donc être redirigé vers la machine hôte. Docker peut se charger de faire la redirection vers un port ne nécessitant pas de privilèges root.

- [1] LXC, <https://linuxcontainers.org>
- [2] OpenVZ, <https://openvz.org>
- [3] Docker, <http://docker.io>
- [4] Go, <https://golang.org>
- [5] Linux control groups, <http://en.wikipedia.org/wiki/Cgroups>
- [6] <https://wiki.archlinux.fr/Chroot>
- [7] UnionFS, A Stackable Unification File System, <http://unionfs.filesystems.org/>
- [8] AuFS, advanced multi layered unification filesystem, <http://aufs.sourceforge.net/>
- [9] carte des périphériques
- [10] Docker Hub, <https://hub.docker.com>
- [11] <http://gl.developpez.com/tutoriel/o...>
- [12] Dockerfile syntax reference, <https://docs.docker.com/reference/b...>
- [13] DevOps, <http://en.wikipedia.org/wiki/DevOps>
- [14] Barrand G. et al., "GAUDI – A software architecture and framework for building LHCb data processing applications", CHEP 2000
- [15] KVM and Docker LXC Benchmarking with OpenStack, <http://bodenr.blogspot.ch/2014/05/k...>
- [16] Felter W. et al., "An Updated Performance Comparison of Virtual Machines and Linux Containers", RC25482, 2014
- [17] App Container Specification and Tooling, <https://github.com/appc/spec>
- [18] rkt, an App Container runtime for Linux, <https://github.com/coreos/rkt>
- [19] The Open Container Initiative, <https://www.opencontainers.org>



n°32
Novembre
2015

La lettre IN2P3 Informatique

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



SC'15 : le CC-IN2P3 et l'Idris représentent le CNRS auprès des grands acteurs du calcul intensif et du big data



© CC-IN2P3 / CNRS

Du 16 au 19 novembre, le CC-IN2P3 a présenté pour la 2e fois un stand commun avec l'Idris à la conférence SuperComputing qui s'est tenue à Austin au Texas (Etats-Unis). Pour la 27e année consécutive, SuperComputing, vitrine majeure du calcul haute performance et du Big Data, a été le rendez-vous incontournable des acteurs industriels et académiques du secteur de l'informatique de pointe pendant une semaine de conférences, de workshops et de présentations. Près de 12 000 visiteurs se sont déplacés pour assister à cet événement qui se veut aussi le plus grand salon au monde où sont présentées les dernières innovations et applications du domaine.

Sur le stand du CNRS, était présentée l'infrastructure du CC-IN2P3 et de l'Idris, en mettant l'accent sur la complémentarité de leur offre HPC et HTC (pour High Throughput Computing). Un focus sur le traitement des données pour le LSST était également là pour rappeler le rôle primordial joué par le CC-IN2P3 dans cette expérience dont les données seront traitées à part égale entre le centre lyonnais et le National Center for Supercomputing Applications (Illinois, Etats-Unis). Au terme des dix ans du programme, ce sont 37 milliards d'objets qui seront stockés, soit un volume total de 500 pétaoctets de données.

L'Idris avait lui choisi de mettre en avant le projet E-Biothon, une plateforme en réseau hébergée au centre parisien et destinée à accélérer et faire progresser la recherche biomédicale. Constituée de plus de 16 000 cœurs de calcul pour une

puissance de 56 téraflopps, cette plateforme permet d'aborder le traitement des données complexes de la biologie d'aujourd'hui afin de mettre au point les logiciels applicatifs de demain. Cette plateforme est issue de la collaboration de plusieurs partenaires : le CNRS, IBM, Inria, l'Institut Français de Bioinformatique et SysFera.

Au-delà de cette présentation, ce stand a également été une fois de plus le point de rencontre et de discussion des différents acteurs du calcul au CNRS. SuperComputing est aussi l'occasion pour le CNRS d'être présent à un niveau international aux côtés d'autres organismes français importants, tels que le CEA ou Inria, avec lesquels un 'French Tour' est organisé, dans le but de renforcer la cohésion et la visibilité de l'ensemble des stands français.

Tout au long de la semaine, plusieurs points forts sont par ailleurs venus rythmer la venue des visiteurs. Le CC-IN2P3 avait misé sur la réalité virtuelle pour mettre en valeur son activité. Une immersion dans le système solaire via un Oculus Rift était proposé, prétexte à présenter l'expérience LSST. Des cardboards permettant de visualiser les vidéos sphériques du centre et du LSST étaient distribuées aux visiteurs les plus concernés. En outre, la société UNIVA, fournisseur du logiciel de gestion de batch du CC-IN2P3, était partenaire du stand et offrait un Oculus au gagnant d'un tirage au sort. Enfin, un Gyropode a été remporté sur le stand par un des participants du French Tour.

Enfin, citons rapidement quelques exemples des grandes tendances dévoilées ou confirmées lors du salon : les technologies liées aux accélérateurs de calcul (GPU, FPGA et Xeon-Phi) qui ont fait l'objet de plusieurs présentations. IBM a par exemple présenté ses neuroprocesseurs (IBM syNAPSE) qui s'inspirent du fonctionnement des neurones pour traiter l'information d'une façon nouvelle et rapide. En ce qui concerne le réseau, on a pu constater que les lignes à 100 Gb/s semblent désormais s'être généralisées. En matière d'infrastructure, on voit apparaître des solutions de refroidissement direct liquide à double phase (liquid cooling). Enfin, on a une nouvelle fois annoncé la mort des bandes magnétiques censées être remplacées par les disques flash...

Gaëlle SHIFRIN



n°32
Novembre
2015

La lettre IN2P3 Informatique

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



Journées LCG France, 14-16 décembre - Villeurbanne (CC-IN2P3)

Les prochaines journées LCG France se dérouleront au Centre de Calcul de l'IN2P3 du lundi 14 au mercredi 16 décembre de cette année.

Vous êtes tous cordialement conviés à ces journées, et vous pouvez vous y inscrire dès maintenant (jusqu'au 30 novembre maximum, le plus tôt étant le mieux).

L'agenda et le formulaire d'inscription sont accessibles *ici*.

L'agenda étant en construction pour le moment, vous pouvez en profiter pour vous manifester auprès du comité d'organisation si vous désirez prendre la parole pour y présenter un sujet d'intérêt pour notre communauté.

Notez que les retours de site Tier-1/-2 /-3 sont toujours appréciés par les collègues des autres sites et les gens impliqués de près dans les expériences LHC.

Vous avez également la possibilité de suggérer des sujets d'échanges pour la session "Qui Veut Parler De Quoi", via l'URL ci-dessus : n'hésitez pas à y proposer ce qui vous semble intéressant de discuter entre collègues pendant cette session informelle.



JRES 2015, du 8 au 11 décembre - Montpellier

Les 11èmes Journées Réseau se tiendront au Corum de Montpellier du 8 au 11 décembre. Plus d'informations sur *le site de l'événement*.

Notons l'importante contribution de l'IN2P3 :

Mardi

- FG-CLOUD : CLOUD COMMUNAUTAIRE DISTRIBUÉ À VOCATION SCIENTIFIQUE
par Jérôme Pansanel - Mohamed Airaj - Catherine Biscarat - Nicolas Clementin - Christine Gondrand - Sébastien Geiger - Vanessa Hamar - Michel Jouvin - Vincent Legoll - Sha Li - Charles Loomis - Matthieu Marquillie - Jean-marc Pierson - Matthieu Puel - Geneviève Romier - François Thiebolt - Andrei Tsaregorodtsev

Mercredi

- DU SMARTPHONE AU FRIGO : LA GRANDE ÉVASION DU MOT DE PASSE
par Serge Borderes
- INTERPRÉTER VOS FLUX AVEC ZNETS2
par Descombes Thierry - Jérôme Fulachier - Ismael

Jeudi

- BUILDING, CONFIGURING, DEPLOYING AND RUNNING DISTRIBUTED, PETA-SCALE DATA ANALYSIS SOFTWARE AT IN2P3
par Fabrice Jammes - Yvan Calas - Fabio Hernandez - Jacek Becla
- FG-IRODS : MUTUALISATION D'EXPERTISE ET INFRASTRUCTURE DISTRIBUÉE POUR UN ACCÈS TRANSPARENT ET HAUTEMENT DISPONIBLE AUX DONNÉES SCIENTIFIQUES
par Jérôme Pansanel - David Benaben - Catherine Biscarat - Yonny Cardenas - Hélène Cordier - Pierre Gay - Benoit Hiroux - Gilles Mathieu - Emmanuel Medernach - Jean-Yves Nief - Geneviève Romier

Vendredi

- CENTRALISATION ET EXPLOITATION DES JOURNAUX SYSTEME
par Jonathan Schaeffer



n°32
Novembre
2015

La lettre IN2P3 Informatique

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



Deux traits d'union de l'informatique IN2P3-IRFU prennent une retraite bien méritée

Remerciements

Beaucoup d'entre nous l'auront compris ; il s'agit de Pierrick Micout et Joseph Le Foll. Tous les deux incarnent le meilleur des relations en informatique entre l'IN2P3 et l'IRFU, l'Institut de recherche sur les lois fondamentales de l'Univers du CEA.



Joseph Le Foll

Pierrick et Joseph, ou, Joseph et Pierrick, ils sont presque indissociables.

Pourtant, l'un est le « joker » de l'autre pour représenter l'IRFU (anciennement DAPNIA) dans toutes les réunions importantes où l'on parle Informatique (Journées Informatique, réunions des expériences du CC, du réseau RI3...). Ils font tous les deux partie des pionniers de l'aventure informatique qui a suivi l'évolution des expériences. Depuis le traitement manuel des données de la chambre à bulles Gargamelle jusqu'à l'aventure des grilles de calcul (de DataGrid à EGI/France Grilles) qui a permis le déploiement de l'infrastructure de traitement des données du LHC aujourd'hui mis en œuvre par WLCG/LCG-France, ils ont tout connu !



Pierrick Micout

C'est avec eux aussi que la collaboration en informatique entre l'IN2P3 et le DAPNIA s'est nouée autour du Centre de Calcul, en 1994 lors de la mutualisation des moyens de calcul des 2 instituts au CC, et ensuite, lors de la connexion à 1Mbit/s avec les États-Unis pour le traitement des données de l'expérience BaBar. Ils ont vu bon nombre de mutations s'opérer et l'informatique devenir porteuse de solutions à l'échelle internationale.

A tous les deux, un grand merci pour votre constance et votre bonne humeur sans faille ! Nous vous souhaitons une longue et belle retraite.

Retrouvez Joseph Le Foll interviewé lors des JI2008 à Obernai : <https://webcast.in2p3.fr/videos-ji0...>