



ÉDITORIAL
Denis Girou

1993 - 2013 :
l'IDRIS a
vingt ans !

Vingt ans au service de ses utilisateurs et de toutes les communautés scientifiques employant la simulation numérique. Ces vingt années, ainsi que le rappelle l'encart que nous consacrons à cet anniversaire, ont vu une prodigieuse évolution technologique qui a démultiplié les puissances disponibles et par là même considérablement accru l'apport de la simulation numérique dans les avancées de la recherche. Cette évolution prodigieuse, dont on sait qu'elle va se poursuivre, va demander à nouveau de fortes adaptations des applications pour tirer substantiellement partie de la puissance potentielle des futures machines. Comme il l'a fait depuis toujours, l'IDRIS continuera, notamment par le biais des formations qu'il propose et de son support applicatif, à aider les communautés scientifiques à se préparer et à s'adapter à ces évolutions nécessaires.

Simultanément, l'IDRIS poursuivra ses actions spécifiques et amplifiera ses participations à divers projets et collaborations. Ceci passe bien sûr par notre engagement continu et essentiel dans les projets européens du domaine du calcul de haute performance, commencé dès 2000 et qui se fait aujourd'hui dans le cadre du *Partnership for Advanced Computing in Europe* (PRACE), dont le cycle suivant, qui débutera en 2015, est en cours de préparation. Mais cela passe aussi par d'autres initiatives et d'autres collaborations. Notre ancrage régional doit se renforcer, en profitant des remarquables opportunités qu'offre la constitution en cours de la future université

Paris Saclay et du grand campus qui va y réunir organismes de recherche, universités et grandes écoles. Dès aujourd'hui, cela se traduira par un renforcement de la participation de nos ingénieurs aux activités de la Maison de la simulation d'Orsay (voir aussi plus loin l'interview de Martial Mancip à propos du projet Digiscope), et par d'autres partenariats locaux comme celui qui vient d'être initié avec le Laboratoire Aimé Cotton, ainsi que l'explique plus loin Olivier Dulieu.

D'autres initiatives en direction de communautés scientifiques particulières ont également été lancées. Pour la communauté de l'environnement, nous allons contribuer au projet ANR Convergence qui vise à faire face, dans le domaine des sciences du climat, aux enjeux du *Big Data* et du passage à l'*Exascale*. Cela concerne également tout spécialement aujourd'hui la communauté des sciences de la vie, jusqu'ici peu présente parmi nos utilisateurs à cause de besoins un peu différents de ceux du calcul de haute performance traditionnel. Ainsi, le projet E-Biothon vient d'être lancé, qui met à disposition une plate-forme constituée d'une partie de la configuration Blue Gene/P exploitée par l'IDRIS de 2008 à 2012, pour accueillir des applications de recherche en biologie et santé. Et puis, toujours dans le domaine des sciences de la vie, c'est le partenariat majeur qui vient d'être établi avec l'Institut français de bioinformatique (IFB) dont l'IDRIS hébergera courant 2014 la plate-forme nationale qui sera mise à la disposition de toute la communauté française de la bioinformatique, tout en devenant le nœud national de l'infrastructure européenne de cette discipline. Jean-François Gibrat expose dans les pages qui suivent le rôle et l'organisation de ce nouvel institut qu'est l'IFB. À côté de nos missions de toujours auprès de la très large communauté de nos utilisateurs, ces nouveaux partenariats, et d'autres qui suivront, nous permettront de renforcer notre double rôle de fournisseur de ressources et de centre d'expertise et d'excellence, afin que les années qui viennent confortent la place que l'IDRIS s'est acquise durant ses vingt premières années d'existence.

sommaire

- 1 • Éditorial
- 2 à 5 • Institut français de bioinformatique
- 6 • Laboratoire Aimé Cotton : interview d'Olivier Dulieu

- 7 • Maison de la simulation : interview de Martial Mancip
- 8 • Projet Convergence
- 8 • Projet E-Biothon
- 8 • Formations de l'IDRIS au 1^{er} trimestre 2014

L'Institut français de bioinformatique :



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



Jean-François Gibrat
directeur de l'IFB, DR1 INRA

Contexte

L'Institut Français de Bioinformatique (IFB) est issu d'une proposition du réseau national des plates-formes (PF) de bioinformatique (ReNaBi) en réponse à l'appel à projets « Infrastructures Nationales en Biologie et Santé » du programme « Investissements d'Avenir ».

Ce projet a été accepté en avril 2012 et doté de 20 M€ sur huit ans. Il vise à créer une infrastructure nationale de services en bioinformatique. En effet, les sciences de la vie produisent maintenant de très grandes quantités de données qu'il faut gérer et analyser, ce qui ne peut se faire que par des techniques informatiques optimisées. Les plates-formes ReNaBi constituant l'IFB sont toutes labellisées par le GIS IBiSA (Infrastructure en Biologie, Santé et Agronomie) qui coordonne les PF de recherche des sciences du vivant. Un des critères déterminants pour être labellisé par IBiSA est d'offrir un service à cette communauté (accès à une infrastructure informatique, mise à disposition d'outils bioinformatiques, de base de données, etc.).

Il est donc important de noter que l'IFB n'est pas directement concernée par la recherche en bioinformatique mais doit fournir un *appui* aux projets de recherche en sciences de la vie. L'objectif général de l'IFB est donc de mettre à disposition de la communauté des sciences de la vie les ressources de base en bioinformatique. Plus particulièrement, ses missions sont :

- de fournir un appui aux programmes de biologie, en termes de soutien aux projets et de formation des biologistes aux outils et méthodes de la bioinformatique ;

- de mettre à disposition une infrastructure informatique dédiée à la gestion et l'analyse des données produites par les sciences de la vie incluant le matériel (stockage et CPU) mais aussi les nombreuses collections de données des sciences

du vivant et les logiciels informatiques pour analyser les données ;

- d'agir comme « intermédiaire » entre la communauté des sciences de la vie et la communauté de la recherche en informatique et bioinformatique.

Structure de l'IFB

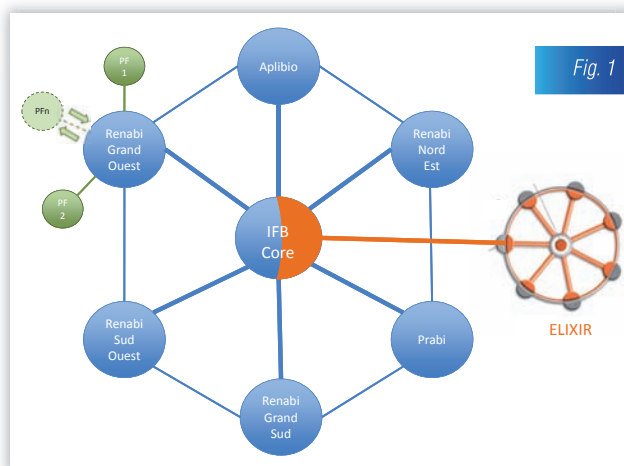


Fig. 1 – Six pôles régionaux : ReNaBi-NE (Lille, Nancy, Reims, Strasbourg), ReNaBi-SO (Bordeaux, Toulouse), ReNaBi-GO (Nantes, Rennes, Roscoff), ReNaBi-GS (Montpellier, Marseille), PRABI (Lyon, Grenoble), APLIBIO (Evry, Jouy, Versailles, Orsay + 5 PF parisiennes). L'IFB sera le nœud français d'ELIXIR.

L'IFB est constitué d'un nœud national, l'IFB-core, et de six pôles régionaux couvrant le territoire national et qui regroupent les 22 PF ReNaBi (cf. Fig. 1). Le nœud national est une unité mixte de service (UMS3601) localisée sur le campus de Gif-sur-Yvette. Elle a été créée en avril 2013 et sa fonction est de prendre en charge les questions administratives et techniques pour l'IFB, entre autres la gestion des 20 M€ alloués.

Pour faire fonctionner cette UMS, il a été demandé aux cinq tutelles (CNRS, CEA, INRA, INRIA, INSERM) dix postes permanents. Les effectifs des pôles régionaux correspondent à environ 100 ETP (équivalent temps plein) et une soixantaine de personnels en CDD.

une infrastructure nationale de service en bioinformatique

Fonctionnement de l'IFB

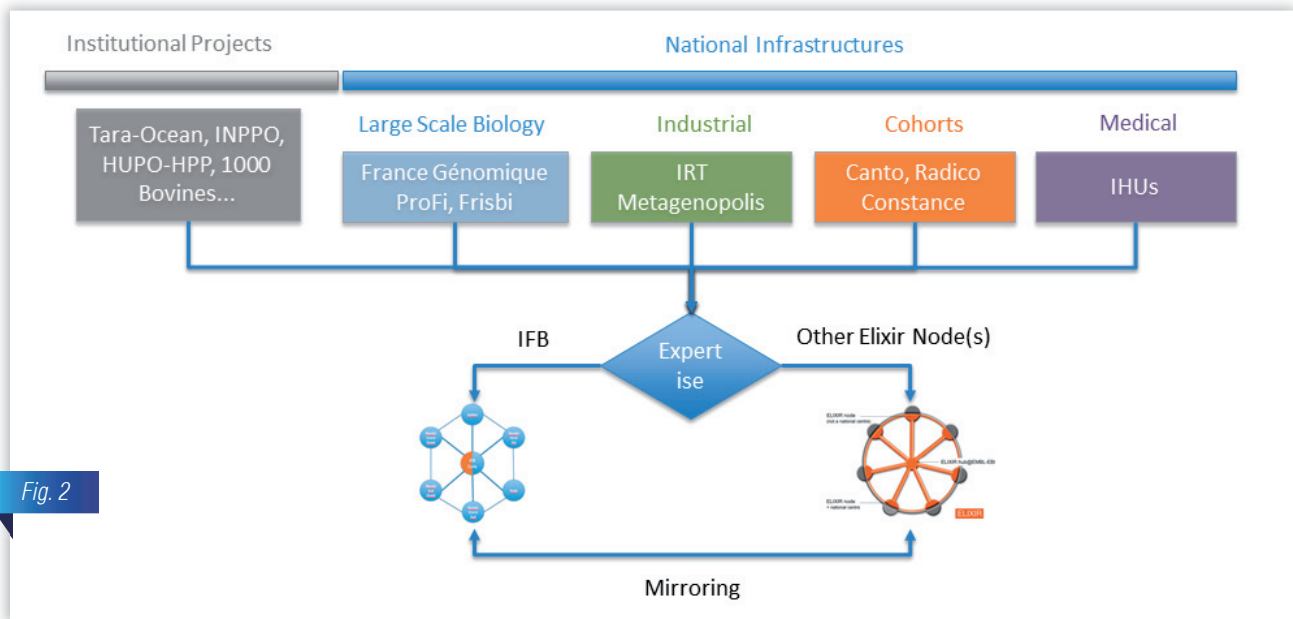


Fig. 2

L'IFB fonctionne en mode projets et fournit un soutien à des projets biologiques, biomédicaux ou technologiques clairement définis. Il y a quatre catégories de projets :

- les projets « institutionnels » à grande échelle ou les projets avec d'autres infrastructures nationales (cf. Fig. 2) ;
- les projets technologiques pour développer des services et intégrer des outils bioinformatiques ;
- les projets de la communauté des sciences de la vie ;
- les services offerts à l'industrie.

Le rôle des PF régionales

Ces dernières :

- offrent un appui et un soutien technique aux projets scientifiques des laboratoires des sciences de la vie (implication dans 160 projets au cours des trois dernières années dont 80 projets d'ANR et 20 européens) ;
- assurent des formations (plus de 1 000 auditeurs par an) ;
- fournissent un libre accès à leurs infrastructures de calcul (plus de 6 000 cœurs et près de 1 Po de stockage ... mais répartis sur plus de 20 PF).

En outre, toutes les PF font partie de laboratoires de recherche ce qui assure ainsi un lien crucial entre la recherche et le service sans lequel ce dernier est amené à périlcliter rapidement.

Le rôle de l'IFB-core

L'IFB-core a pour tâches de :

- fournir un soutien administratif et technique à l'IFB ;
- compiler les indicateurs de l'activité de l'infrastructure pour les tutelles ;
- fournir un point d'entrée clairement identifié à la communauté des sciences de la vie ;
- fournir un accès centralisé aux ressources offertes par les PF IFB (offres d'appui, services, logiciels, bases de données, infrastructures informatiques) ;
- aider à structurer et coordonner les activités des PF ;
- mutualiser les ressources et l'expertise dans les domaines suivants :
 - la mise à disposition d'une infrastructure informatique dédiée aux sciences de la vie ;
 - les formations des utilisateurs ;
 - le soutien aux projets technologiques de R&D.

L'IFB-core et core-IT

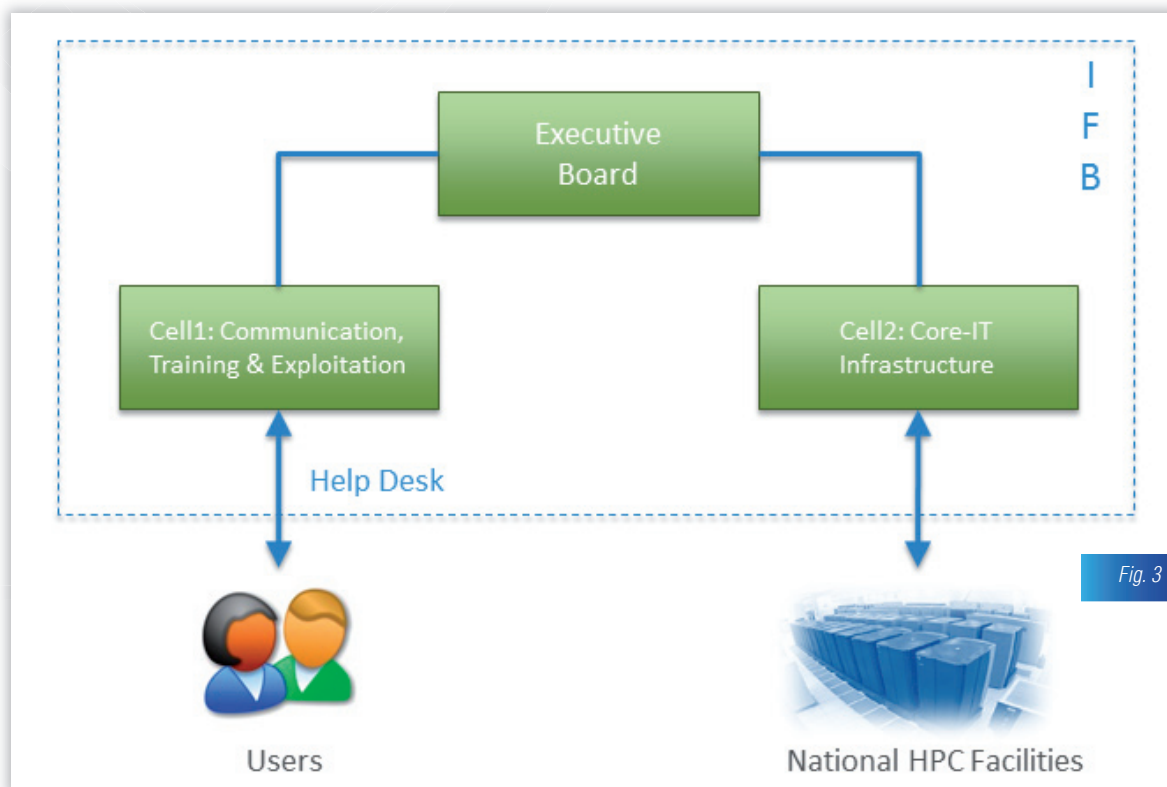


Fig. 3

L'UMS3601 (IFB-core) est localisée dans le bâtiment Imagif sur le centre de Gif-sur-Yvette. Elle est constituée des cellules : « Communication, Formation et Valorisation » et « Infrastructure ». (Fig. 3)

Nous ne décrivons pas plus avant les fonctions de la première cellule. La cellule « Infrastructure », quant à elle, est chargée de mettre en place et de faire fonctionner l'infrastructure nationale de l'IFB (appelée « core-I T ») qui sera hébergée à l'IDRIS. Cette infrastructure est dimensionnée pour atteindre, à terme, 10 000 cœurs et 1 Po de stockage (elle sera en taille sensiblement similaire à la machine ADA de l'IDRIS). La tâche la plus importante de cette cellule sera de mettre en place un Cloud académique dédié à la gestion et au traitement des données des sciences de la vie, s'appuyant, dans un premier temps, sur l'infrastructure nationale de l'IFB.

Dans un second temps, et suivant l'avancée des solutions informatiques, ce Cloud sera étendu aux infrastructures régionales des PF IFB sous la forme d'une fédération de Clouds. Pour mener à bien cette tâche critique, l'IFB sera amené à collaborer étroitement avec l'IDRIS et l'Institut des Grilles et du Cloud du CNRS.

Les caractéristiques de la bioinformatique

Les données produites par les sciences de la vie

La bioinformatique, dans une acception large, est définie comme l'ensemble des méthodes informatiques permettant de gérer et d'analyser les données produites par les sciences du vivant. Cependant, il faut noter qu'historiquement la communauté bioinformatique s'est développée à partir de l'analyse des données de séquences (séquences d'ADN, d'ARN, de protéines), de leurs structures et de leurs interactions et n'inclut pas l'analyse des images qui constituent aussi des données biologiques très importantes. Les technologies actuelles permettent d'obtenir de grandes masses de données sur les systèmes biologiques. En particulier, depuis 2007 les technologies de séquençage sont capables de produire des téraoctets (To) de données par expérience (par exemple un séquenceur HiSeq peut produire plus d'un milliard de lectures



appariées de 2x100 nucléotides, soit 1,6 To rien que pour les données « nues », le double avec les métadonnées de qualité). L'intérêt de toutes ces nouvelles technologies, c'est qu'elles permettent aux biologistes d'adopter une démarche encyclopédique dans l'étude des organismes en leur donnant accès à l'ensemble du génome d'un organisme, à tous les transcrits, toutes les protéines, toutes les interactions, tous les métabolites et les flux, etc. Le revers de cette médaille, c'est que les biologistes sont noyés sous un flot de données hétérogènes et extrêmement volumineuses. Au-delà de la simple gestion de ce flot de données, le rôle central que joue la bioinformatique est d'aider les biologistes à extraire des connaissances biologiques à partir de masses de données « brutes ».

Les caractéristiques des analyses bioinformatiques

Une caractéristique essentielle des analyses bioinformatiques est qu'elles sont, le plus souvent, facilement distribuables. Si on dispose de N processeurs, il suffit de découper les données en N lots qui sont exécutés indépendamment sans que les processeurs aient besoin de communiquer entre eux. Une grande majorité des traitements bioinformatiques ne nécessite donc pas l'utilisation des machines HPC des grands centres de calcul où les processeurs coopèrent grâce à de coûteuses liaisons à large bande passante et faible latence telle qu'InfiniBand.

Une autre caractéristique des traitements bioinformatiques est l'extraordinaire foisonnement des logiciels d'analyse des données et la grande diversité des langages de programmation utilisés. Nous avons ainsi recensé 98 logiciels rien que pour effectuer l'alignement des lectures produites par les séquenceurs sur des génomes de référence, qui ne constitue, pourtant, que la première étape d'un certain nombre d'autres analyses de ces données. La plupart des PF de bioinformatique mettent ainsi à disposition de leurs utilisateurs plusieurs centaines de logiciels différents pour toutes sortes d'analyses. Beaucoup d'analyses sont constituées par un enchaînement de logiciels, les sorties des uns étant les entrées des autres d'où l'intérêt manifesté par les bioinformaticiens pour les « workflows » et autres « pipelines ». En outre, un certain nombre de logiciels de bioinformatique ont des dépendances à d'autres logiciels sous forme d'appels à des bibliothèques. Par exemple, beaucoup de logiciels utilisent Perl qui est un langage souvent employé en bioinformatique et pour lequel il existe des modules spécialisés dans l'analyse de données biologiques (BioPerl). Il existe différentes versions

du langage et des modules BioPerl qui sont censées être compatibles d'une façon ascendante. Cependant, il suffit que différents logiciels d'un pipeline d'analyse nécessitent des versions différentes des modules BioPerl ou Perl pour que l'installation du pipeline sur un nouveau système ne devienne assez pénible. C'est sans doute une des raisons pour laquelle l'utilisation des grilles a eu du mal à s'imposer dans la communauté bioinformatique, l'autre raison étant l'utilisation de grosses quantités de données mises à jour régulièrement et devant être accédées en mode fichier pour des questions de compatibilité avec l'ensemble de ces outils logiciels historiques. Nous espérons que le Cloud permettra de répondre plus facilement au premier problème en fournissant des machines virtuelles prêtes à l'emploi et encapsulant tous les logiciels d'un pipeline ainsi que les différents modules et bibliothèques requis pour un type d'analyse donné (« appliances » bioinformatiques).

Une dernière caractéristique des traitements bioinformatiques est leur étroite dépendance à l'évolution très rapide des technologies de production de données. Les technologies de séquençage apparues dans les années 2007 ont commencé à générer de grandes quantités de données mais avec des lectures très courtes (seulement 35 nucléotides pour les premières versions). À l'époque, les logiciels existants ne savaient traiter efficacement que des lectures plus longues et n'étaient pas du tout adaptés au grand nombre de lectures générées. Très rapidement, la communauté recherche en bioinformatique a étudié de nouveaux algorithmes qui ont été encapsulés dans différents logiciels (par exemple, pour l'assemblage des lectures, passant d'algorithmes basés sur la programmation dynamique à ceux basés sur les graphes de de Bruijn). Les technologies de séquençage ont évolué depuis 2007, produisant toujours autant de données mais augmentant aussi continuellement la longueur des lectures. Du coup, les algorithmes développés pour les lectures très courtes ne sont plus efficaces et il va falloir, une fois de plus, développer de nouveaux algorithmes. Toutes ces caractéristiques ont fait que la communauté bioinformatique n'a pas eu tendance à utiliser les grands centres de calcul du GENCI, hors cas particuliers, comme les simulations de dynamique moléculaire, et a plutôt développé ses propres infrastructures en local. Néanmoins, la multiplicité des infrastructures locales a un coût qui devient de plus en plus élevé et un des rôles de l'IFB sera d'essayer de mutualiser les ressources informatiques de la communauté dans la mesure du possible.

Interview d'Olivier Dulieu, directeur-adjoint du Laboratoire Aimé Cotton

par Denis Girou, Thierry Goldmann et Geneviève Morvan

L'IDRIS vient d'accueillir dans ses locaux, pour une période de quelques années, une vingtaine de chercheurs du Laboratoire Aimé Cotton (LAC).

– Dites-nous tout d'abord quelles sont les thématiques scientifiques de votre laboratoire et quels sont plus précisément les domaines de recherche des personnels qui viennent de s'installer à l'IDRIS ?

OD – Les thématiques scientifiques du LAC recouvrent les domaines de la physique atomique, moléculaire, l'optique, l'information quantique, la nanophysique et la biophysique ainsi que le bio médical, en particulier l'aide au handicap. Les personnels qui viennent à l'IDRIS sont pour la plupart des théoriciens en physique atomique et moléculaire et en information quantique, ainsi qu'un groupe de quatre expérimentateurs en nanophysique.

– Quelles sont les raisons qui ont conduit à ce partenariat avec l'IDRIS ?

OD – Le point de départ a été les travaux de rénovation importants entrepris dans le bâtiment du LAC dans le cadre du projet de « Centre de Physique Matière et Rayonnement » (CPMR) prenant place dans le contexte de la nouvelle université Paris-Saclay. Cela nécessitait de trouver temporairement des locaux à l'extérieur du LAC, tout en permettant aux expérimentateurs de demeurer proches de leurs équipements, d'où l'idée de s'adresser en priorité aux théoriciens du laboratoire. M^{me} Debischop, directrice de la délégation régionale Île-de-France Sud a alors mentionné l'opportunité de s'installer au troisième étage de l'IDRIS. Cette opération a représenté un intérêt bien compris des deux parties car l'accueil d'équipes scientifiques dans ses locaux rentre dans les missions de l'IDRIS et s'est pratiqué à diverses époques. C'est une belle aventure à tenter.

– Au-delà des motifs pratiques qui ont donc poussé à cette installation à l'IDRIS, y voyez-vous aussi une opportunité pour développer l'utilisation de la simulation numérique dans vos thématiques de recherche ? Que représente-t-elle dans vos activités d'aujourd'hui et comment entrevoyez-vous la place qu'elle devrait prendre pour vous dans les années à venir ?

OD – Effectivement, le point évident qui permet d'envisager cette coopération scientifique, c'est l'utilisation de la simulation



Une partie des membres du LAC dans les locaux rénovés du 3^e étage de l'IDRIS

numérique et du calcul intensif dans nos thématiques. Cela faisait longtemps que nous faisons déjà appel à des petits clusters de calcul (notamment celui de la Fédération de Recherche LUMière MATière sur le campus d'Orsay), suffisamment puissants aujourd'hui pour traiter toute une catégorie de problèmes physiques. Par ailleurs, nous avons recruté récemment des jeunes chercheurs et c'était l'opportunité de définir de nouveaux projets scientifiques, avec une composante plus marquée concernant la simulation numérique sur des calculateurs à haute performance comme ceux de l'IDRIS. Nous les avons encouragés, nous les encourageons, à suivre les formations aux techniques du calcul intensif dispensées à l'IDRIS.

Nous avons d'ores et déjà un projet scientifique de dynamique collisionnelle avec un code numérique parallélisé qui avait atteint les limites des possibilités du cluster de LUMAT, et qui sera le premier code porté sur l'un des supercalculateurs de l'IDRIS dans les prochaines semaines.

Dans les années à venir, nul doute que la simulation numérique, qui pour nous s'inscrit essentiellement dans les sujets traités par le comité thématique Chimie quantique et modélisation moléculaire (CT8), prendra une place croissante, aussi bien en chimie quantique (calculs de structures électroniques) qu'en dynamique collisionnelle (matière froide et ultra-froide).

– Pour conclure, comment abordez-vous cette nouvelle phase dans la vie de votre laboratoire ?

OD – Nous abordons cette nouvelle phase avec beaucoup d'optimisme et de confiance ! Ce partenariat est une richesse supplémentaire pour le LAC et l'IDRIS. Nous avons été accueillis par l'IDRIS dans des conditions tout à fait remarquables et la DR4 a apporté tout son concours technique et logistique à notre installation, avec l'appui de l'INS2I et l'INP. Cela va être pour nous une nouvelle façon de travailler, avec un nouveau rythme à trouver entre la localisation à l'IDRIS et les équipes restées au LAC.

Interview de Martial Mancip, ingénieur Visualisation CNRS - Maison de la Simulation

par Thierry Goldmann

– Présentez-nous les activités de la Maison de la Simulation.

MM – La Maison de la Simulation est un laboratoire regroupant cinq partenaires (CEA, CNRS, INRIA, université d'Orsay et université de Versailles – St-Quentin) qui a pour objectif d'accompagner et de soutenir les communautés scientifiques afin de tirer le meilleur parti des supercalculateurs, en particulier ceux déployés dans le cadre de GENCI et du projet européen PRACE. Elle favorise notamment l'émergence en France d'une communauté du calcul intensif et développe les synergies entre chercheurs et ingénieurs de différentes disciplines nécessaires pour concrétiser les avancées scientifiques importantes attendues du calcul haute performance.

– La Maison de la Simulation installe dans ses locaux un mur d'images à très haute résolution dans le cadre de l'Equipex DIGISCOPE. Pouvez-vous nous décrire cet équipement ?

MM – Le projet DIGISCOPE est une infrastructure de visualisation de haute performance pour l'interaction collaborative avec des données et des calculs massifs. L'objectif est de déployer un ensemble de plateformes de type murs d'images interactifs et salles de réalité virtuelle immersive interconnectées par des réseaux à hauts débits et des moyens de télé-présence permettant la collaboration à distance. Le mur d'images de la Maison de la Simulation sera composé de 16 écrans Barco 52 pouces permettant une taille d'image maximale de 4,30 m de large sur 2,40 m de haut et d'une résolution de plus de 33 millions de pixels. Ce sera le plus grand mur en Europe utilisant cette technologie de cubes de rétro projection LED HD. Associé à un cluster graphique Hewlett Packard composé de 48 cœurs, 2 To de mémoire et 4 nœuds de rendu dotés de 8 GPU Kepler 5000 Nvidia, ce mur d'images Très Haute Résolution permettra d'utiliser toutes les applications de visualisation scientifique avec un rendu en « temps réel » 3D.

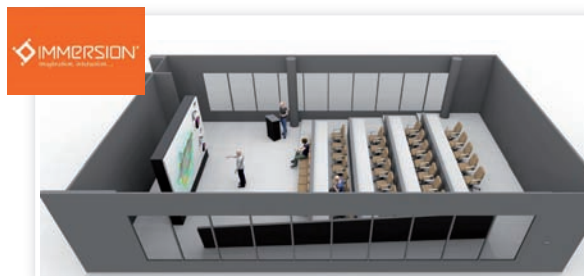
– Ce mur d'images est destiné à quels usages ?

MM – Il faut rappeler que c'est un mur THR (Très Haute Résolution) permettant l'équivalent de l'affichage de plus de 16 écrans à la norme Full HD (2 millions de points par écran) en vision monoscopique et ce, sur plus de 10 m² ! Cela permet donc de visualiser des objets graphiques complexes en ne sacrifiant pas le niveau de détail, ou des maillages très volumineux, des systèmes moléculaires organisés, des simulations cosmologiques denses. On est donc loin de la réduction de vision imposée par l'usage d'un écran traditionnel sur lequel il est très difficile par exemple de visualiser des champs de vecteurs dans des maillages à haute résolution (2048, 4096 au cube), où on ne peut que s'intéresser à une petite portion du domaine. De plus, la possibilité d'utiliser la

vision stéréoscopique sur autant de pixels permet de ne pas réduire le volume de données calculées.

– Comment un utilisateur de l'IDRIS pourra visualiser ses résultats sur votre mur d'images ?

MM – Tout d'abord, nous attachons une très grande importance au confort de travail. Il faut que l'utilisateur puisse visualiser ses résultats en conservant ses habitudes en terme d'interface homme-machine. Il pourra utiliser le même logiciel de visualisation que sur son cluster local. Il est essentiel qu'il puisse interagir avec les données produites et non rejouer à l'infini une séquence vidéo où il ne serait possible que de changer le coefficient de grossissement ou le point de vue de l'observateur. Il pourra ainsi utiliser Paraview ou Visit (entre autres) en profitant de la puissance de calcul et de rendu du cluster en amont du mur d'images. Visualiser une molécule complexe ou un amas galactique sur 10 m² de surface en immersion permet une approche des phénomènes radicalement différente ! Ensuite, à moyen terme, nous pouvons imaginer, à l'aide du réseau à haut débit disponible entre l'IDRIS et la Maison de la Simulation, qu'un utilisateur lance une simulation sur plusieurs milliers de cœurs de votre machine massivement parallèle Turing et, en utilisant Visit en mode client/serveur, visualise en temps réel les résultats sur le mur d'images tout en interagissant en permanence sur les données. Et ceci, de mémoire à mémoire, sans écrire un seul octet sur disque.



Vue d'artiste de la salle Mandelbrot

– Quelles évolutions voyez-vous à moyen et long terme à apporter à votre mur d'images ?

MM – Il est important d'encourager et de développer la vision stéréoscopique d'objets complexes afin d'améliorer la compréhension des phénomènes qu'on visualise, de développer des interfaces homme-machine utilisant les techniques d'interaction sans fil, l'interconnexion entre deux murs d'images où, grâce aux outils de télé-présence déployés dans DIGISCOPE, on peut imaginer deux façons de représenter un même phénomène et donc, en temps réel, d'interagir sur l'un ou l'autre des murs d'images pour avancer dans son domaine scientifique. Nous entrerons dans un nouveau paradigme, celui de la visualisation collaborative, indispensable pour la résolution des simulations à grande échelle.

ACTUALITÉS

Projet Convergence

Le projet Convergence a démarré en octobre 2013, pour une durée de quatre ans.

C'est un projet ANR « Modèles numériques » mené par l'Institut Pierre Simon Laplace (IPSL), avec pour partenaires le CERFACS, le Centre de recherche de la météorologie nationale, l'IDRIS et la Maison de la simulation. Il est axé sur les thématiques du *Big Data* et de la préparation à l'*Exascale* pour les sciences du climat : en effet, l'ambition d'utiliser des résolutions plus fines nécessite un passage à l'échelle de tous les modèles utilisés ainsi que des moyens décuplés pour analyser les données générées. L'objectif principal de ce projet est de préparer le prochain exercice CMIP (*Coupled Model Intercomparison Project*) pour le *World Climate Research Programme* (WCRP), pour pouvoir ensuite en réaliser les simulations ainsi que la mise à disposition simultanée des données produites, pour l'ensemble de la communauté scientifique de cette discipline. Pour plus d'information, vous pouvez consulter le site Web du projet : <http://convergence.ipsl.fr/>

Projet E-Biothon



Le CNRS, IBM, l'Inria, l'Institut français de bioinformatique et la start-up innovante SysFera déploient E-Biothon, une plateforme expérimentale mise à la disposition des chercheurs en bioinformatique pour accélérer et faire progresser la recherche en biologie, santé et environnement. Cette plateforme, hébergée à l'IDRIS, est constituée d'une partie de notre ancienne configuration Blue Gene/P et sera accessible au travers d'un portail applicatif dédié. Cela permettra aux chercheurs de ces disciplines d'effectuer des traitements sur les données complexes de la biologie d'aujourd'hui afin de mettre au point les logiciels applicatifs de demain et d'accélérer la recherche en biologie et en santé, en particulier en génomique et en protéomique. Pour plus d'information, consulter le communiqué de presse du CNRS du 19 novembre 2013 : <http://www2.cnrs.fr/presse/communiqu/3313.htm>

CALENDRIER DES FORMATIONS IDRIS PROGRAMMÉES AU PREMIER TRIMESTRE 2014

| TITRE DE LA FORMATION | DATE DE DÉBUT | DURÉE |
|----------------------------------|---------------|-----------|
| Fortran 95-1 | 28/01/2014 | 3 jours |
| MPI | 03/02/2014 | 4 jours |
| Utilisation Blue Gene/Q | 10/02/2014 | 2 jours |
| OpenMP | 18/03/2014 | 2 jours |
| Fortran 95-2 | 25/03/2014 | 3,5 jours |
| Programmation hybride MPI/OpenMP | 07/04/2014 | 4 jours |

Ces dates vous sont communiquées à titre d'information et sont susceptibles d'être mises à jour. Pour une information récente et plus complète, n'hésitez pas à consulter le serveur Web des cours de l'IDRIS : <https://cours.idris.fr>. Vous y trouverez le catalogue complet des formations et pourrez vous préinscrire aux sessions annoncées.

Nous vous rappelons que les formations IDRIS sont gratuites pour les personnes appartenant au CNRS ou à une université. Elles sont aussi accessibles au personnel d'entreprises publiques ou privées via CNRS Formation Entreprises : les conditions d'inscription sont alors consultables sur le site Web : <http://cnrsformation.cnrs.fr/>