DECEMBRE 2013

HighPerformanceComputing

Le magazine référence du calcul intensif et du Big Data

Volume I | Numéro 9

France

0x09

/DECOUVRIR

SIMULER LA TURBULENCE

Revue de détail des modèles classiques et hybrides

/VERBATIM

JEAN-FRANCOIS LAVIGNON

Une vision globale, au-delà de Bull et de l'ETP4HPC

/HPC_LABS

MAITRISER OPENACC 2.0 (1/3)

Pratiquez les nouvelles fonctions de gestion de données

/L'ETAT_DE_L'ART

LES BIBLIOTHEQUES DE CALCUL

Panorama et perspectives d'un écosystème de plus en plus riche

HPC ET MEDECINE : DES AVANCEES DECISIVES

Recherche, pratiques de pointe, approches prédictives : le calcul au service de l'*ars medica*...

WWW.HPCMAGAZINE.FR

Abonnez-vous
Consultez nos archives
Découvrez nos contenus exclusifs

Une publication HPC Media

EDITO

DECEMBRE 2013

Format Desktop / Tablettes

Jingle bells

Cher Père Noël, toute l'équipe du magazine et moi-même t'adressons ici notre liste de souhaits. Nous nous en remettons à ton infinie bonté - sachant que nous avons été on ne peut plus sages. Le premier d'entre eux, tu t'en doutes, c'est qu'à l'avenir le temps nous soit donné de paraître avec un peu moins de retard ! Si tu pouvais user de ton influence pour que les dates des grandes conférences ne coïncident pas avec les dates de bouclage des mensuels, ce serait cool.

Ensuite, nous serions très heureux que tu apportes à nos éditions Europe et Americas le même accueil chaleureux que celui que l'édition France a reçu cette année. C'est pour nous une grande aventure, pleine de défis, mais les encouragements que nous avons reçus nous portent.

Enfin, bien que tu aies déjà gratifié l'équipe de deux nouveaux collaborateurs de grande valeur - Stéphane Bihan et Eric Tenin - d'autres talents encore ne seront pas de trop. A ce propos, merci également pour les nombreuses contributions qui ont permis au magazine, mois après mois, de se distinguer par la qualité de son contenu.

Ne sois donc pas surpris si, de notre côté aussi, nous avons pensé à toi (regarde page 40). Et n'oublie pas que le soir du 24, un bon Pommard t'attendra à la rédaction.

Frédéric MILLIOT
Directeur de la rédaction
frederic@hpcmagazine.com

CARRI Systems

Architecte technologique depuis 1992

intel

Découvrez notre gamme **XLR dédiée aux accélérateurs sur ports PCI Express**

Les nouveaux serveurs de calcul XLR de CARRI Systems sont désormais compatibles avec les coprocesseurs Intel® Xeon® Phi™ 5110P.

HighServer XLR4i

Appuyez vos innovations de pointe avec le traitement hautement parallèle du coprocesseur Intel® Xeon Phi™. Chaque puce peut traiter plus d'un téraFLOP en performances de pointe à double précision, les meilleures performances de traitement parallèle par watt de tous les processeurs Intel® Xeon®

intel inside™ XEON PHI™

Coprocesseur Intel® Xeon® Phi™

HighServer XLR4i
133548

- Intel® Xeon® E5-2620 à 2 Ghz
- 32 Go DDR3 1600 MHz ECC REG
- 2 x 1 To RAID Edition (7 200 tr/min, 64 Mo de cache)
- 1 x Intel® Xeon® Phi™ 5110P avec 8 Go GDDR5 @ 5.0 GT/s
- Chipset vidéo intégré
- 2 x Gigabit Ethernet
- Rack 2U
- 2 x alimentations redondantes 1620 W
- Linux Centos®
- Garantie 3 ans sur site pièces et main d'oeuvre (hors accessoires)
- France Métropolitaine uniquement (hors îles)

À partir de 5 990 euros HT

Découvrez nos solutions XLR sur

www.carri.com

intel Technology Provider Platinum 2013

intel inside™ XEON®

intel inside™ XEON PHI™

intel inside™ SOLID-STATE DRIVE

Ultrabook, Celeron, Celeron Inside, Core Inside, Intel, Logo Intel, Intel Atom, Intel Atom Inside, Intel Core, Intel Inside, Logo Intel Inside, Intel vPro, Itanium, Itanium Inside, Pentium, Pentium Inside, vPro Inside, Xeon, Xeon Phi, et Xeon Inside sont des marques de commerce d'Intel Corporation aux Etats-Unis et dans d'autres pays.
Intel, le Logo Intel, le Logo Intel Inside, Xeon et Xeon Phi sont des marques de commerce d'Intel Corporation aux États-Unis et dans d'autres pays.

AU SOMMAIRE

CE MOIS-CI

/actus
06 L'essentiel
ROMEO fait des jaloux, Intel dévoile Xeon Phi 2.0, NVIDIA sa K40 et AMD ses APU, Dongarra honoré, LINPACK bientôt remplacé et + encore...

/en_couverture
16 HPC et médecine
La démocratisation du calcul intensif et des technologies Big Data touche aussi le médical. Où qu'on se tourne, les progrès sont exponentiels...

/verbatim
30 J.-F. Lavignon, Bull/ETP4HPC
A l'ordre du jour : l'ETP4HPC "dans le dur", les enjeux de l'indépendance technologique européenne et des API pour parvenir à l'Exascale

/decouvrir
40 Simuler la turbulence
Modèles classiques, modèles hybrides, des solutions existent aujourd'hui pour simuler les phénomènes de turbulence les plus complexes

/l'etat_de_l'art
47 Les bibliothèques de calcul
Enjeux économiques et sociétaux, gains en productivité, l'écosystème actuel des bibliothèques scientifiques favorise l'innovation collective

/hpc_labs
55 Maîtriser OpenACC 2.0
Ce premier article d'une série de trois se consacre aux nouvelles fonctions de gestion de données, un domaine qui méritait vraiment d'être amélioré...

© HPC Media 2013.

Ont collaboré à ce numéro :

Philippe d'Anfray
Stéphane Bihan
Stéphane Chauveau
Amaury de Cizancourt
Alain Clapaud
Gilles Eggenspieler
Michel Kern
Thiên-Hiêp Lê
Frédéric Milliot
Fabien Nicolas
Eric Tenin

+33 (0) 9 50 96 64 00

REDACTION
Suggestions, commentaires, propositions d'articles...
laredaction@hpcmagazine.fr

ABONNEMENTS
Un problème avec votre abonnement ?
abonnements@hpcmagazine.fr

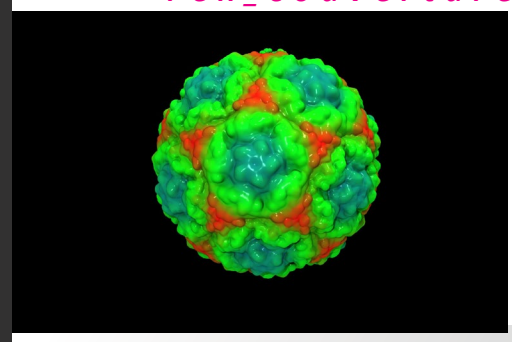
PUBLICITE
Annoncez dans le seul média intégré dédié à l'univers HPC.
pub@hpcmagazine.fr

/actus

Distingué
N°5 au Green500 (et 151ème au Top500) - belle récompense pour ROMEO !

Page 6

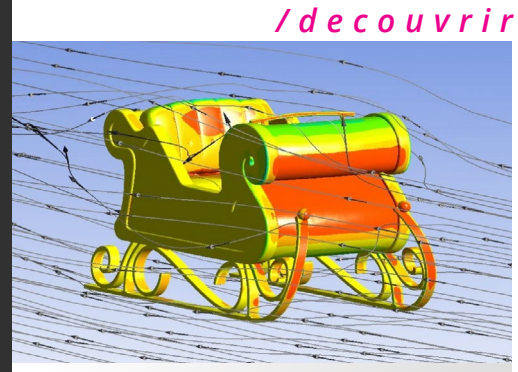
/en_couverture



Décrypté
Virtualisé au niveau atomique, le rhinovirus devient une cible facile

Page 17

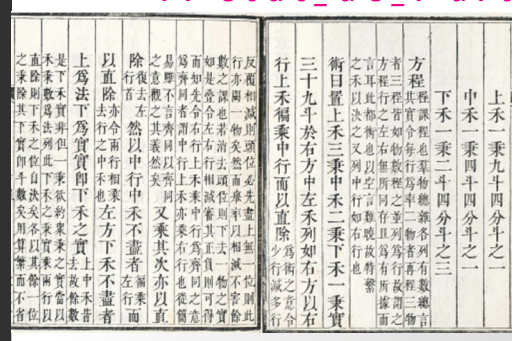
/decouvrir



Turbulent
Les nouveaux modèles hybrides autorisent toutes les simulations industrielles

Page 40

/l'état_de_l'art



Contributif
Voilà 21 siècles que le partage du savoir mathématique fait avancer le monde

Page 45

HighPerformanceComputing
HPC MEDIA
11, rue du Mont Valérien
92210 St-Cloud - France

PNY®

SERVEURS HPC de calcul par GPU

Accélérez les procédures de calcul scientifique et technique avec les GPU NVIDIA Tesla.
Les serveurs TYAN et ASUS de PNY, associés aux GPUs Tesla de dernière génération, constituent l'architecture la plus rapide et la plus efficace au monde.

Plus d'informations sur produits, visitez www.pny.eu/tesla et le configurateur PNY
http://pny.quadrok-selector.com/

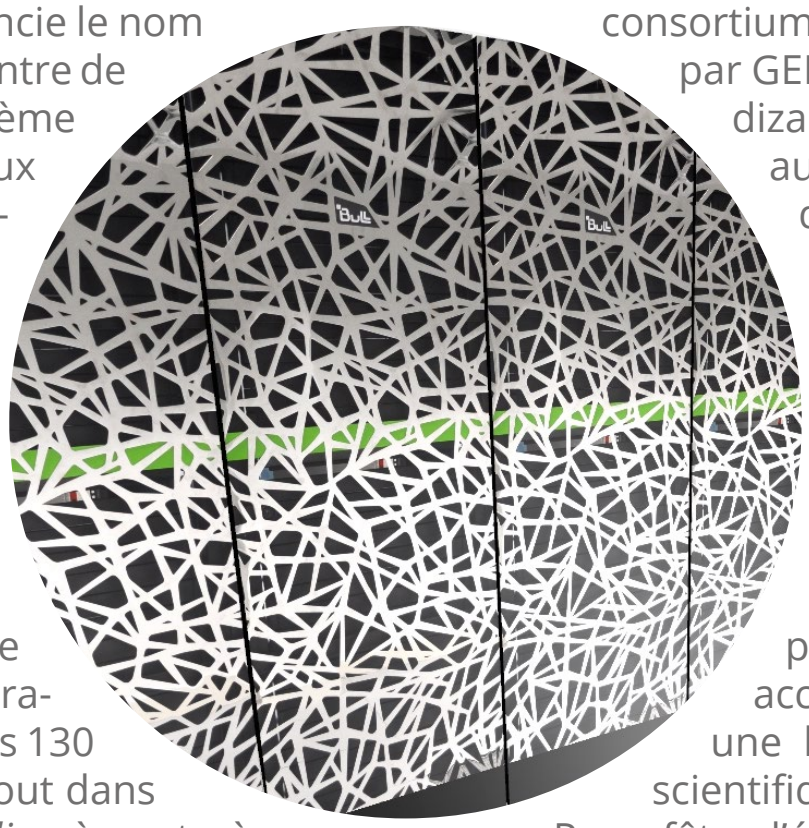
@PNYProFR

ASUS® TYAN

/ACTUS

ROMEO au pinacle !

Number 5 au Green500, voilà le classement que vient d'obtenir l'Université de Reims Champagne-Ardenne pour le cluster du centre de calcul ROMEO. Calculateur le plus éco de France, Roméo (le *camel casing* différencie le nom de la machine du nom du centre de calcul) est également le 151ème plus puissant au monde, deux "bonnes notes" qui récompensent assurément le travail accompli de longue date par Michaël Krajecki et Arnaud Renard, en partenariat avec Bull et NVIDIA pour ce projet.



Roméo, dans le détail, c'est 250 Tflops de puissance crête issus de 260 accélérateurs Kepler K20X logés dans 130 serveurs Bullx R421 E3, le tout dans une infrastructure *free-cooling* à porte à eau. Après l'inauguration officielle prévue pour le 17 décembre, cette artillerie lourde sera rendue disponible à la fois aux chercheurs mais aussi aux industriels de la région, avec un encouragement à l'utilisation par les PME-PMI. Financé par le Fonds européen de développement régional (FEDER) et par les Investissements d'Avenir du gouvernement, soutenu par la Région et la ville de Reims, le mésocentre ROMEO fait partie du consortium Equip@meso coordonné par GENCI. Il compte ainsi parmi la dizaine de sites Tier-2 ouverts aux entreprises dans le but de favoriser l'innovation et la compétitivité via la simulation numérique.

Outre ses ressources de calcul, ROMEO offrira à ses utilisateurs académiques et privés un espace de stockage sécurisé, une pile logicielle complète et un accompagnement expert dans une large palette de domaines scientifiques et techniques avancés.

Pour fêter l'événement, ROMEO a d'ailleurs organisé "**Exceed the limits!**", une dotation de 5 millions d'heures offertes aux chercheurs européens pour un certain nombre de projets sélectionnés par le comité scientifique du centre. Pour tout cela, un grand Bravo à l'équipe !

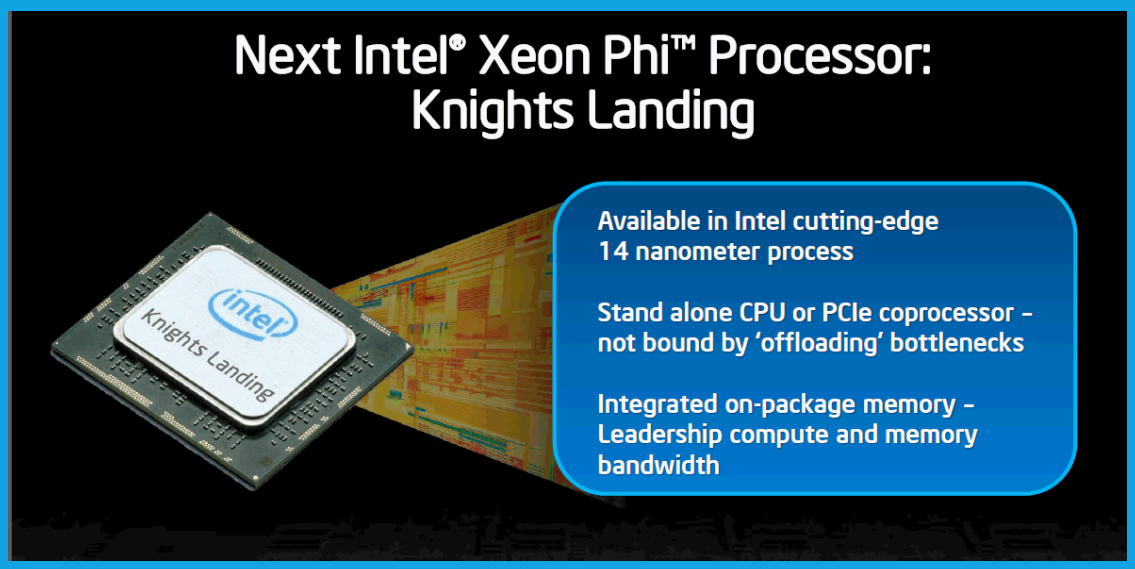
CAPS à louer

Votre code ne profite pas assez des accélérateurs ou des CPU multicores ? Vous souhaitez évaluer le gain de performance potentiel de vos applications sur les coprocesseurs Intel Xeon Phi, les GPU NVIDIA Kepler ou les APU AMD ? C'est pour répondre à ces besoins si particuliers que CAPS Entreprise, dont l'expertise est unanimement reconnue au niveau international, vient d'annoncer le lancement de sa nouvelle entité. CAPS Engineering, tel est son nom, va se consacrer au conseil et à l'accompagnement en parallélisation de code source. Toujours basées à Rennes, les équipes pourront se déployer partout dans le monde et offrir à leurs clients le même niveau de compétence que celui qui fait des compilateurs et outils de développement CAPS des références dans leur domaines. Bon vent au team CAPS pour cette nouvelle aventure !

/actus

XEON PHI : INTEL SE MET À TABLE

Un an près l'annonce officielle de Xeon Phi en version "Knights Corner", Intel a profité de SC pour organiser une table ronde dédiée au futur de son accélérateur x86. Concrètement, il s'agissait pour Intel de préciser publiquement ses intentions à long terme, et de rappeler que l'objectif de Xeon Phi est d'assurer une transition totale vers le multicore et le manycore, c'est-à-dire de permettre aux applicatifs le passage à un très grand nombre de cœurs dans un contexte de programmation x86 conventionnelle.



*Aujourd'hui, le principal problème du calcul hétérogène est le chargement des données sur les accélérateurs. Intel souhaite apporter y une réponse durable avec l'architecture mémoire de Knights Landing : un mécanisme de cache situé sur le cœur lui-même et des mémoires à bande passante élevée (*in-package high bandwidth memory*) qui, connectées à de la DDR4 en grande quantité, doivent assurer une capacité et une bande passante suffisante. L'utilisation de ce cache pour contrôler le placement mémoire et exploiter les cœurs supplémentaires devrait permettre à Knights Landing, toujours selon les ingénieurs Intel, d'exécuter les applications nativement plutôt qu'en mode accéléré, et ainsi d'éviter les transferts de données.*

Reste qu'on en sait désormais plus sur ce qui caractérisera concrètement Phi 2.0 : gravé en 14 nm et principalement disponible en format socket, il embarquera 72 cœurs Airmont (Intel Atom) avec instructions out-of-order *vectorielles sur 512 bits (AVX-512), 8 à 16 Go de mémoire directement sur l'accélérateur, 6 canaux d'accès à la future mémoire DDR4 (jusqu'à 384 Go à 115 Go/s !) et 36 lignes PCIe gen3. Question performances, 3 Tflops en double précision sont annoncés, soit 3X les performances de Knights Corner dans un TDP compris entre 160 et 200 W. La version carte fille se limitera à deux canaux de DDR4 pour un maximum de 64 Go de RAM. Une version "F" devrait également être disponible, avec des interconnexions optiques (le "F" est pour Fabric) et un TDP de 15 W supérieur. Quant à la disponibilité de ces petites bêtes, rien n'a changé : selon la version, il faudra attendre de mi à fin 2015 pour pouvoir s'en déclarer propriétaire.*

NEC SX-ACE : 16 Tflops par armoire vectorielle

Enterré le vectoriel ? Sûrement pas ! NEC vient en effet de présenter SX-ACE, successeur du valeureux SX-9, qui améliore la puissance de calcul d'un facteur 10 et réduit la surface au sol par 5. Cette machine est équipée d'un processeur propriétaire contenant quatre cœurs vectoriels pouvant réaliser chacun jusqu'à 64 Gflops et dotés d'une bande passante de 64 Go/s, ce qui en fait tout simplement les cœurs les plus performants au monde !

Côté dimensionnement, une armoire SX-ACE embarque jusqu'à 64 modules contenant chacun un processeur quadri-cœur, l'ensemble délivrant pas moins de 16 Tflops. Il est possible d'en connecter jusqu'à huit, la performance crête maximale atteignant alors 131 Tflops. Si techniquement SX-ACE représente un certain état de l'art pour la parallélisation massive, reste à convaincre les utilisateurs de sortir de l'univers x86 et d'adopter de nouvelles méthodes de programmation via les compilateurs et bibliothèques NEC.

/actus

K40, CUDA 6 : NVIDIA SE FOCALISE SUR LA PERFORMANCE

C'était la grosse annonce de SC'13 pour le spécialiste de l'accélération parallèle par GPU. Grâce à l'amélioration continue des *processes* de fabrication, la carte Tesla K40 succède à la Tesla K20X au sommet de l'architecture Kepler, avec notamment un doublement de la mémoire (12 Go de GDDR5), un nombre de cœurs actifs porté à 2880 (au lieu de 2688), une fréquence légèrement augmentée (745 MHz au lieu de 735) et une connexion PCIe gen3 doublée - le tout dans la même enveloppe thermique. Résultats pour les performances crête : 4,29 Tflops en simple précision, 1,43 Tflops en double précision.

Selon les quelques *benches* présentés par le constructeur, les *speedups* mesurables se situent entre 20 et 40 % - les résultats les plus élevés étant obtenus sur les applications les plus gourmandes en mémoire (CFD, sismique...). Enfin, NVIDIA annonce également pour cet accélérateur un mode "GPU boost" destiné à augmenter temporairement la fréquence de tous les cœurs simultanément (jusqu'à 875 MHz) et contrôlé par le développeur. Voilà qui demande à être testé aux HPC Labs...

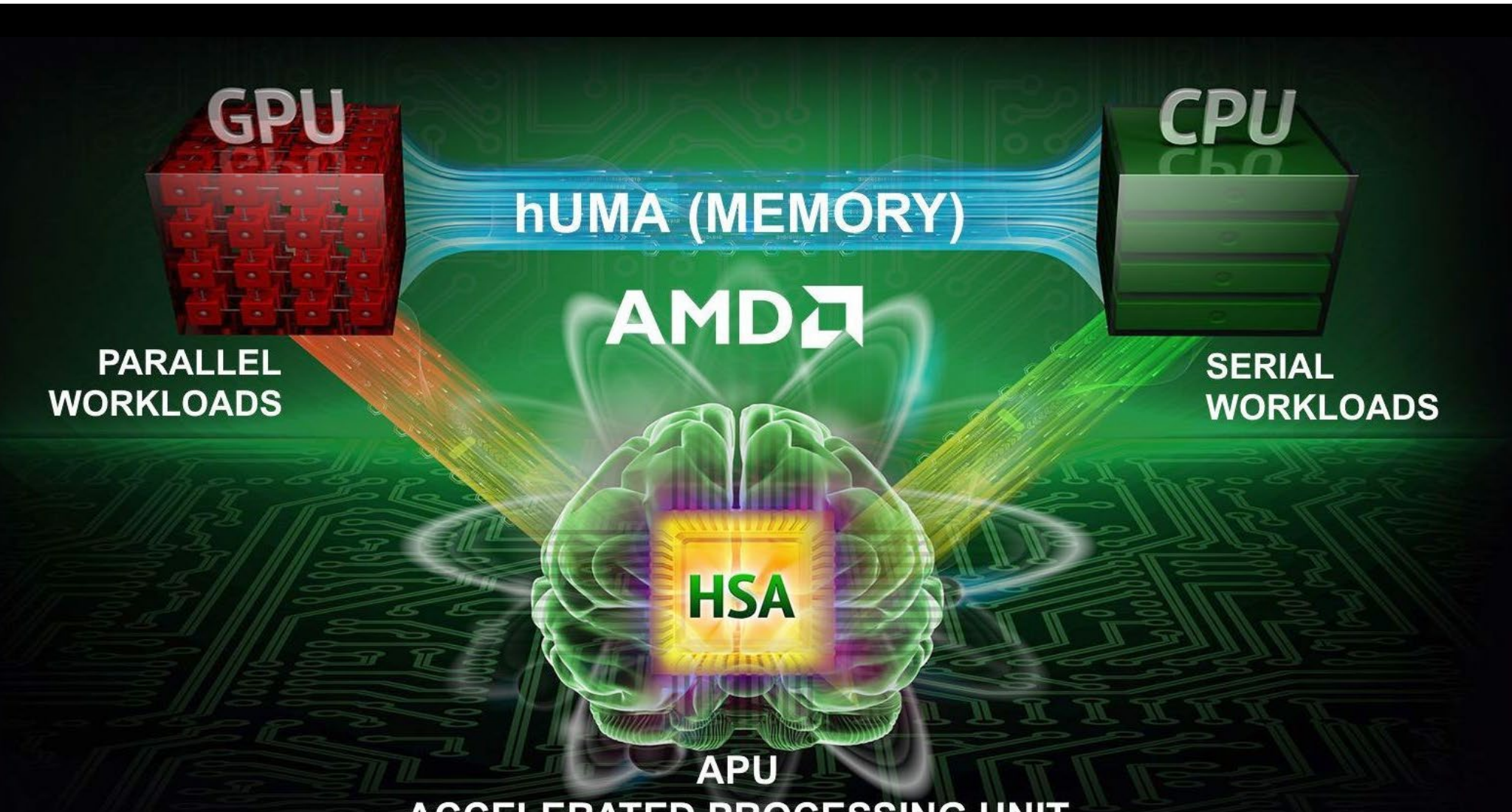
CUDA 6 est dans la place

Fiançailles avant l'union sacrée ? La nouvelle version de CUDA, également annoncée à SC, unifie logiquement la mémoire, pour la plus grande joie du développeur qui n'a plus à sa charge la gestion explicite des transferts de données entre CPU et GPU. Rien à voir, donc, avec l'union physique des deux espaces, qui supprimerait définitivement *THE* goulet d'étranglement, mais celle-ci reste en ligne de mire. Indépendamment de l'Unified Memory, CUDA 6 apporte également des bibliothèques *drop-in* qui remplacent automatiquement les versions CPU de BLAS et FFT par leur équivalent GPU, ainsi que l'optimisation multiGPU de ces bibliothèques (jusqu'à 8 GPU par nœud). Une release importante, donc, sur laquelle nous reviendrons en détail très prochainement.

Amazon Web Services adopte les GPU

Leader mondial dans le domaine du HPC *as a Service*, AWS étoffe son portefeuille d'offres de calcul à distance avec de nouvelles instances Amazon EC2 (Elastic Compute Cloud) baptisées NVIDIA G2 ("NVIDIA GRID GPU"). Equipées de processeurs graphiques NVIDIA Kepler, ces instances s'ajoutent à la trentaine de services déjà proposés par Amazon aux entreprises et organisations déjà utilisatrices - globalement ou ponctuellement - de calcul et de visualisation à la demande. Point important, NVIDIA G2 inclut un puissant moteur d'encodage H.264 dédié à la compression graphique en haute résolution. Grâce à celui-ci, on peut envisager de travailler à distance sur virtuellement tous types de terminaux, notamment mobiles, quelle que soit l'application hébergée. Le fait qu'un poids lourd comme AWS s'y mette aussi devrait accélérer l'adoption du "télé-HPC". En tous cas aux Etats-Unis, où les problématiques d'hébergement des données sur le sol national ne sont pas les mêmes que dans le reste du monde...

AMD reprend l'initiative



Déjà présents sur les PC de bureau et les consoles de jeu, les APU AMD (qui combinent sur le même *die* CPU et GPU) entrent au royaume des serveurs. Avec l'architecture Kaveri, disponible début 2014, on peut donc s'attendre à de réels gains en densité et en efficacité énergétique. Pour préparer les applicatifs à cette nouvelle donne technique, AMD et ses partenaires ont également travaillé sur les outils de programmation correspondants. On attend notamment de nouvelles versions de Java Virtual Machine, des compilateurs OpenMP GCC et PGI Accelerator, des bibliothèques OpenCL clMath d'AMD et ArrayFire 2.0 d'AccelerEyes, ainsi que de l'outil maison CodeXL pour un débogage et un profilage dignes de ce nom.

Au delà de cette annonce, la vision officielle d'AMD est d'unifier l'ensemble des plateformes matérielles de l'industrie, à commencer par les siennes, autour de l'architecture HSA (Heterogeneous System Architecture) sur laquelle seront basées les prochaines générations d'APU. Réunissant d'autres intervenants de poids comme ARM, Ti, Qualcomm et Samsung, l'objectif de la fondation HSA est de proposer une seule et même abstraction matérielle pour tous les besoins d'hétérogénéité. Voilà qui devrait singulièrement faciliter le travail des développeurs : une seule architecture, c'est en effet un seul code exécutable dans une large variété d'environnements…

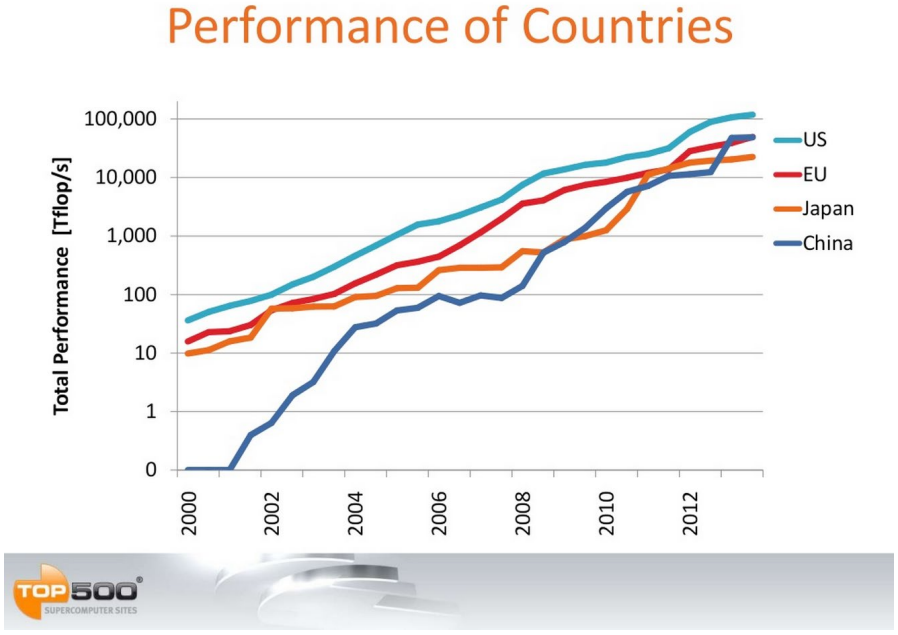
Pour concrétiser cet effort d'unification tant attendu de la part de la communauté, un nouveau SDK basé sur OpenCL a été annoncé pendant l'AMD Developer Summit 2013. Il reprend toutes les fonctionnalités de l'APP SDK 2.9 en incluant des outils multimédia pour le traitement des vidéos, la bibliothèque clMath et, bien sûr, CodeXL. Voilà qui confirme plus que jamais l'approche Open Source privilégiée par AMD pour la programmation de ses processeurs - une approche qu'on a probablement décrié trop vite.

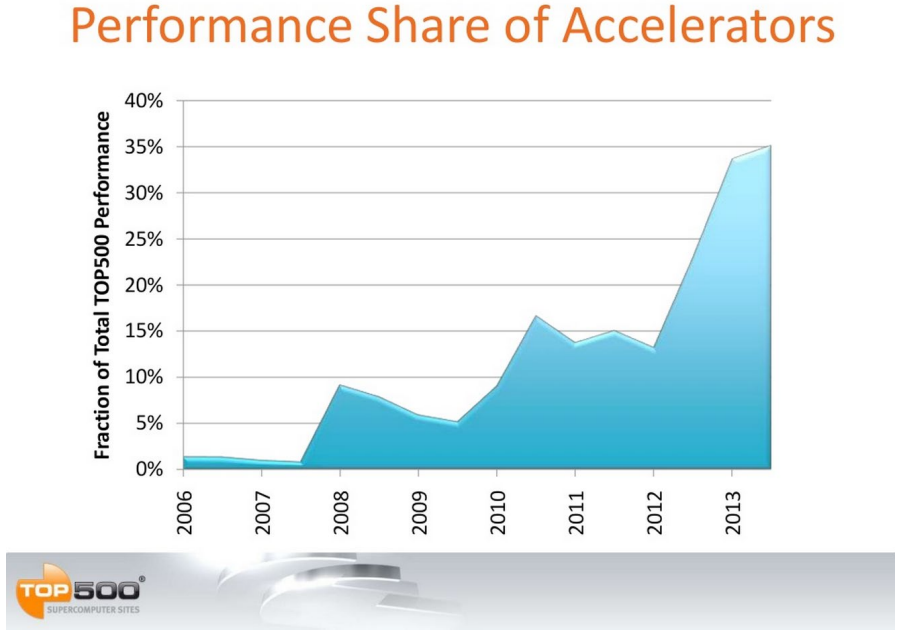
Ca bouge (un peu) dans les classements mondiaux

Fatalité intangible, compétition sans excitation ? Le premier des calculateurs mondiaux ce semestre est à nouveau le chinois Tianhe-2 avec ses 33 Pflops, loin - très loin - devant Titan, pour l'heure encore vaisseau amiral de l'ORNL américain, et ses 17 Pflops. Notons l'entrée surprise dans le top 10, à la 6ème place pour être précis, de Piz Daint (5,6 Pflops), le système CRAY XC30 du CSCS Suisse avec une efficacité énergétique de 2.6 Gflops / W.

compte pour 63 entrées, deux de moins qu'en juin. En Europe, le Royaume-Uni, la France et l'Allemagne peuvent revendiquer respectivement 23, 22 et 20 des systèmes les plus puissants.

Côté Green500, carton plein pour NVIDIA. Il y a 6 mois, seuls deux clusters à base de GPU Tesla figuraient au top 10 des machines vertes. Désormais, tous les calculateurs de ce top 10 embarquent les accélérateurs de la marque. Pour





Rien de bien nouveau côté Top500. En revanche, le Green500 monte en puissance. Logique...

Ce n'est pas au sommet que le Top500 de novembre 2013 se renouvelle, c'est dans les couches intermédiaires. D'abord, on remarque de plus en plus de nouveaux systèmes pétaflopiques - 31 au total. Par ailleurs, le nombre de machines utilisant les technologies accélératrices, 53, n'a pas varié depuis juin : 38 sont choisi l'architecture NVIDIA Kepler, 13 l'architecture Intel MIC (Xeon Phi) tandis que deux calculateurs sont boostés par des GPU AMD. Point intéressant, 75 % des systèmes classés sont au minimum octocores (95 % sont au minimum hexacores).

D'un point de vue géographique, les Etats-Unis dominent toujours largement le paysage HPC mondial avec plus de la moitié des systèmes répertoriés, suivis de l'Asie (115 systèmes) et de l'Europe (102). En Asie, la Chine uniquement

comparaison, lors du classement de juin 2012, les vingt premières machines étaient des IBM BlueGene/Q. La machine la plus efficace cette année s'appelle Tsubame-KFC (rien à voir avec la friture de nuggets - encore que...). Installée au Tokyo Institute of Technology, elle affiche une efficacité énergétique de 4,5 Gflops par Watt et se positionne devant Wilkes (Université de Cambridge), machine européenne la plus efficace à 3,6 Gflops / W. Petit clin d'œil également à Roméo, la toute nouvelle machine de l'URCA, qui se hisse d'emblée en cinquième position (voir notre actu à ce sujet p. 6). Il faut également remarquer une réelle montée en puissance des *green systems*. Si les précédents classements montraient des puissances crête relativement modestes, le top10 comprend maintenant 3 calculateurs pétaflopiques. Ca avance, moins vite qu'on le souhaiterait, mais ça avance...

/actus

Jack Dongarra officiellement excellent

Scientifique bien connu de la communauté pour ses travaux sur les outils logiciels et les bibliothèques de calcul, le Pr Jack Dongarra vient d'être récompensé du Ken Kennedy Award pour son "*leadership dans la conception et la promotion de standards pour les logiciels mathématiques utilisés dans la résolution des problèmes numériques en calcul haute performance*".

Le prix Kennedy a, rappelons-le, été créé en 2009 en hommage à feu Ken Kennedy, fondateur de l'université Rice et expert HPC lui aussi internationalement renommé. Sponsorisée par l'ACM et la Computer Society, cette distinction récompense chaque année des contributions exceptionnelles à la programmabilité et la productivité du calcul informatique. Tout le profil de Jack exactement car, si son nom est associé à la création de LINPACK, benchmark qui a longtemps fait référence pour l'évaluation des performances des calculateurs, il a également contribué de façon décisive à l'implémentation de nombreuses bibliothèques Open source. On citera entre autres BLAS, LAPACK, Netlib, MPI, OpenMP-MPI, ATLAS, PLASMA et MAGMA - rien que du *world-class*...

Ce qu'il y a de particulier, avec Jack Dongarra, c'est que ses innombrables occupations professorales ne l'empêchent pas de parcourir le monde, inlassablement, pour apporter généreusement ses lumières personnelles à qui les sollicite. Avec la curiosité et la créativité qui l'animent, il travaille aujourd'hui sur un certain nombre de problématiques logicielles liées à l'Exascale, et plus particulièrement sur **PaRSEC**, un ordonnanceur de tâches pour les architectures hétérogènes manycores à mémoire distribuée. Nous associons humblement nos félicitations à celles du prestigieux Jury.

Micron transforme la mémoire en processeur parallèle

Von Neumann pourrait bien se retourner dans sa tombe ! Plutôt que d'améliorer sans cesse la bande passante des mémoires qui alimentent nos processeurs conventionnels, pourquoi ne pas tout simplement inverser le problème et utiliser la mémoire non plus uniquement comme simple espace de stockage mais comme élément de *processing* actif ? C'est ce qu'a fait Micron, également inventeur de l'Hybrid Memory Cube, avec Automata Processor. A quoi ressemble ce concept pour le moins révolutionnaire, qui pourrait concurrencer coprocesseurs et accélérateurs en calcul massivement parallèle?

Imaginez un FPGA contenant une grille de centaines voire de milliers d'unités de calcul programmables qui, contrairement aux GPU exécutant la même opération en parallèle sur des Go de données, parallélisent des opérations différentes. Voilà qui, sur le papier en tous cas, est pour le moins séduisant. La programmation d'un tel composant, c'est-à-dire la transformation d'une application donnée en une machine dédiée, n'aura cependant rien d'un jeu d'enfant. Pour la faciliter, Micron travaille à un SDK qui devrait être disponible avant l'été. Affaire à suivre, clairement !

L'exascale au sommet du Mont-Blanc

SuperComputing a permis au groupe de recherche en architectures hétérogènes de BSC, dirigé par Alex Ramirez, de dévoiler un troisième prototype de ce à quoi pourraient ressembler les futures machines exaflopiques qui devront allier puissance et faible consommation. Qui dit efficacité énergétique dit évidemment GPU et CPU basse consommation, ceux-là mêmes qui motorisent nos mobiles préférés. C'est donc sur des processeurs de ce type que le troisième prototype de Mont-Blanc est basé : des ARM Cortex-A15 combinés à des GPU Mali-T604 sur le même *die*. Pour mémoire, le précédent prototype, Petraforca, dévoilé plus tôt cette année, contenait un CPU quadri-core Cortex-A9, deux fois moins puissant que le dual-core Cortex-A15, et utilisait un coprocesseur NVIDIA Kepler K20.

Ce nouveau cluster se distingue également par le système de châssis et de rack des lames Bullx B505. C'est dans cet écrin que se logent les 1620 cœurs ARM et les 810 accélérateurs graphiques, qui ensemble offrent une puissance crête de 26 Tflops. On compte six châssis contenant chacun neuf lames embarquant quinze cartes de calcul Mont-Blanc. Une lame délivre environ 485 Gflops pour 200 W - un chiffre à comparer au 1,4 Tflops servis par Kepler K40 pour une enveloppe de 235 W. C'est probablement avec la prochaine itération, qui inaugurera les CPU ARM v8 en 64 bits, que l'efficacité énergétique de ce design prendra tout son sens. Le projet venant de recevoir 8,1 M€ de la part de la Commission Européenne, sa pérennité est assurée pour au moins deux années supplémentaire.

LINPACK bientôt remisé ?

Le débat sur l'utilité de Linpack est-il en passe d'être clos ? De plus en plus nombreux, les critiques du vieux benchmark commencent à se réjouir. Conçu à la fin des années 70 pour mesurer la seule capacité de calcul en virgule flottante, LINPACK n'est plus du tout représentatif du comportement de nos applications sur les machines contemporaines, notamment en ce qui concerne le réseau d'interconnexions. D'où l'intérêt qu'a suscité Jack Dongarra, autorité incontestée sur l'affaire puisqu'il en est le créateur, en annonçant pendant SC'13 travailler avec son équipe sur un nouveau protocole de test baptisé HPCG (High Performance Conjugate Gradient). L'objectif, évidemment, est de refléter plus exactement les simulations exécutées aujourd'hui sur les gros clusters. Une première version alpha du benchmark va très bientôt être disponible pour validation sur un large spectre de machines.

Dongarra prévient cependant que même si cette alpha a déjà permis de sortir des chiffres exploitables, le moment n'est pas encore venu de remplacer LINPACK. Quelques trimestres de rodage vont encore être nécessaires pour que HPCG acquière la reconnaissance de la communauté et, ce faisant, supplante son ancêtre.

Toujours est-il que les premiers résultats issus de HPCG révèlent des écarts de performance assez spectaculaires, d'un facteur 40 à 50 par rapport à LINPACK ! Ce qui signifie que le train-train du classement Top500 a de grandes chances d'être bouleversé. D'aucuns y verront sans doute l'occasion de privilégier enfin des architectures plus en rapport avec les besoins applicatifs réels...

Densité record chez Supermicro !

A l'occasion de l'annonce de NVIDIA K40, Supermicro a présenté une nouvelle gamme de serveurs, lames et stations de travail, avec pour la majorité des modèles une densité inédite - jusqu'à 20 accélérateurs dans un châssis 7U !

Parmi les nombreuses références au catalogue du constructeur, citons le SuperServer qui offre pas moins de huit GPU et deux processeurs Intel Xeon E5-2600 v2 Ivy Bridge, le tout en 4U. Compte tenu de la dissipation thermique nécessaire sur un tel design, deux systèmes de refroidissement séparés y sont intégrés.

Supermicro propose également un large éventail de serveurs lames SuperBlade qui sur 1U, 2U ou 3U supporteront de trois à six GPU au total. Côté stations de travail, la gamme s'enrichit d'une SuperWorkstation, bête de course à cinq GPU intégrés qui devrait faire fureur en Alaska et au Groenland. On pourra y loger deux Xeon E5-2600 (v1 ou v2), jusqu'à 1 To de DDR3 ECC, huit unités de stockage 3,5 pouces et un large éventail de cartes d'extension (4 PCIe 3.0 x16, 2 PCIe 3.0 x8 et 1 PCIe 2.0 x4). L'ensemble est pourvu d'une alimentation redondante Platinum de 1620 W.

Allinea analyse la performance parallèle

En complément du débogueur DDT et du profiler MAP, Allinea a profité de SuperComputing pour dévoiler Allinea Performance Reports, un nouvel outil d'analyse de la performance d'exécution des applications HPC hautement parallèles qui ne nécessite ni instrumentation, ni recompilation. Son principal objectif : déterminer si les codes sont adaptés à la plateforme matérielle sur laquelle ils doivent être exécutés. Voilà qui rend la suite d'outils de développement toujours plus intégrée et plus productive !

Pour faire bonne mesure, l'éditeur a également annoncé le support de nouvelles cibles, notamment le coprocesseur Intel Xeon Phi et ARM v7 au travers la plateforme NVIDIA CUDA 5.5. En toute logique, CUDA 6 devrait suivre sans tarder...

Les HPC Days 2014

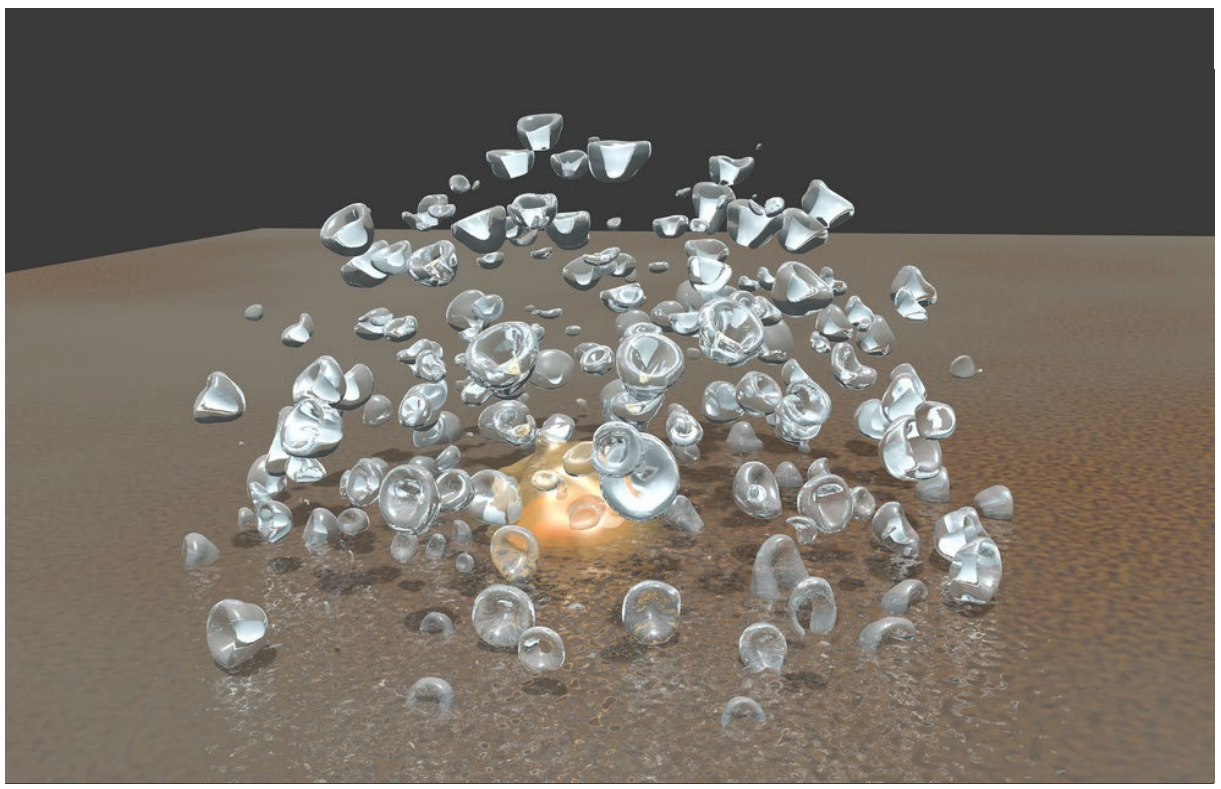
Du HPC, rien que du HPC ! Voilà ce que propose l'UNISTRA pour les deux jours de formation qu'elle organise, en parallèle avec un séminaire scientifique, en ses locaux strasbourgeois les 22 et 23 janvier prochains.

Au menu : La programmation CUDA (session assurée par NVIDIA), OpenACC (présenté par NVIDIA), Le débogage d'applications parallèle avec TotalView (session finement intitulée "Un bug, combien ça coûte" et assurée par NEC / Rogue Wave), L'optimisation de codes (session assurée par le pôle HPC de l'université) et Programmation sur Xeon Phi (Pôle HPC as well - jeu de mot !).

Les inscriptions démarrent mi-décembre. Venez nombreux !

Des bulles pour le prix Gordon Bell

Dans notre précédent numéro, nous vous présentions les six projets finalistes du prix Gordon Bell 2013, dont quatre avaient utilisé Titan pour réaliser leurs simulations. Il faut croire que les statistiques n'étaient pas de la partie car la prestigieuse récompense a finalement été attribuée à l'équipe dirigée par Michael Norman (SDSC de San Diego + EPFL de Zurich) pour ses travaux sur la cavitation.



En résumé, treize trillions de cellules réparties sur 6,4 millions de threads ont été utilisées pour résoudre un phénomène unique de disparition des bulles de gaz dans un liquide soumis à une fluctuation de pression. Réalisées sur Sequoia, le calculateur IBM BlueGene/Q du laboratoire Lawrence Livermore, cette simulation représente une amélioration de la performance d'environ 150X par rapport à l'état de l'art.

Les retombées de cette recherche vont permettre d'améliorer la conception des moteurs à combustion et des turbines hydrauliques. Ils devraient également impacter très positivement un certain nombre d'approches thérapeutiques du cancer. Bravo à l'équipe, donc, mais aussi aux autres finalistes dont la sélection favorisera sans doute la carrière.

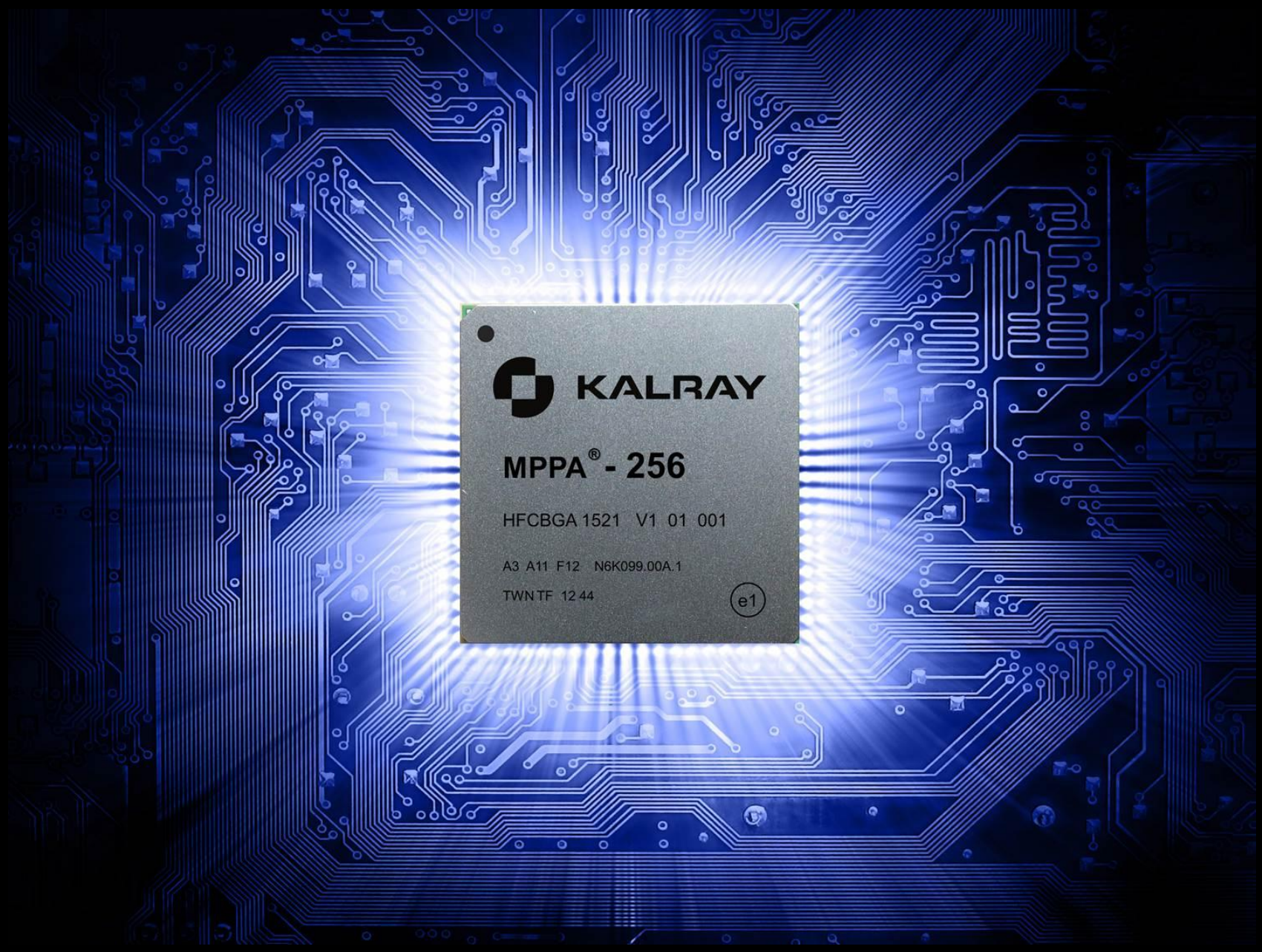
Tous

nos articles,
nos dossiers,
nos chroniques,
nos codes sources
et bien plus encore...

www.hpcmagazine.fr

KALRAY

AGILE PERFORMANCE



Kalray MPPA®-256

256 cores, 25 GFlops/Watt

World first massively parallel, scalable, MIMD processor optimized for high performance computing and power consumption

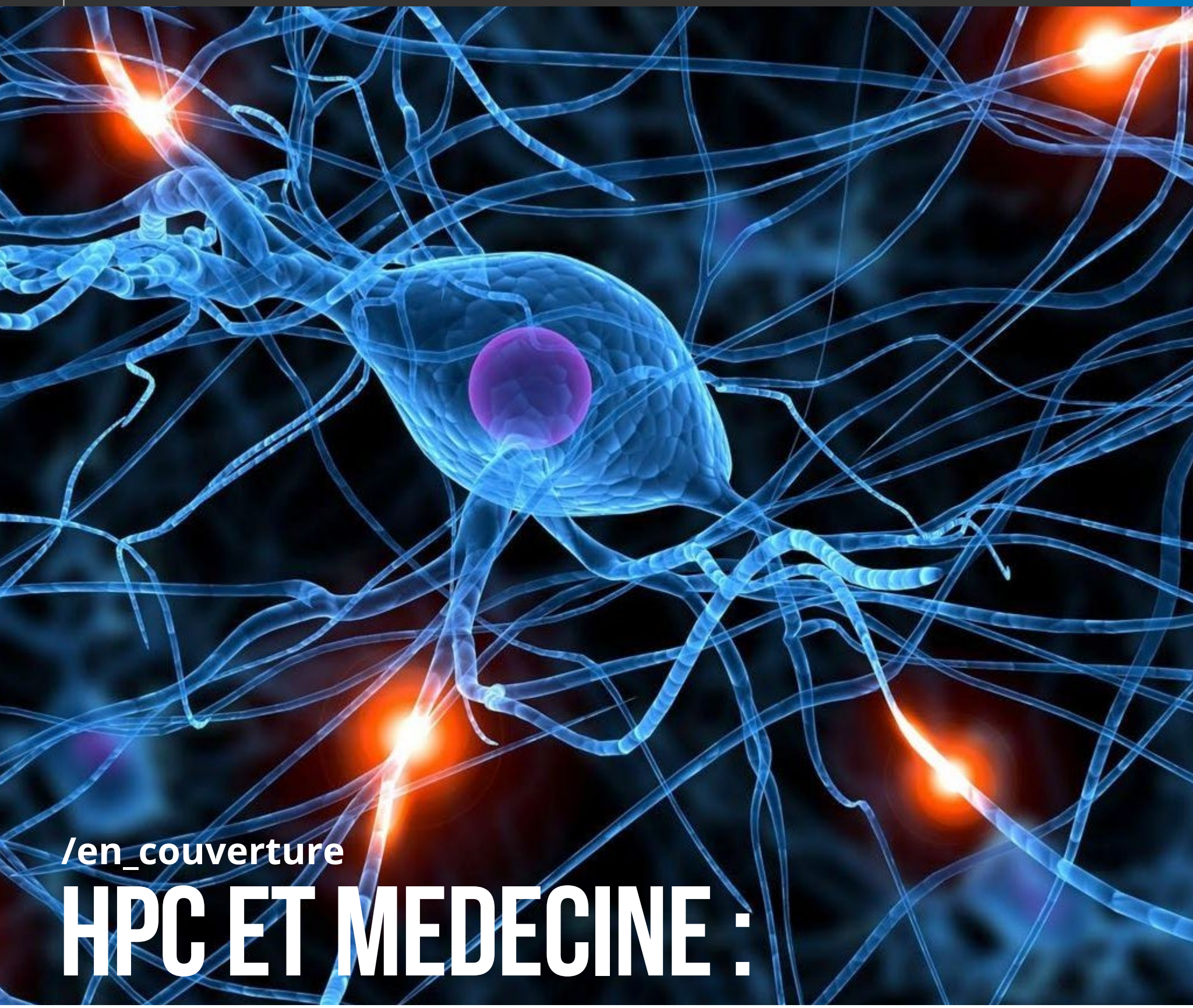
www.kalray.eu

Development workstation
MPPA DEVELOPER

SDK C/C++/FORTRAN, debugger and trace
MPPA ACCESSCORE

Reference boards
MPPA BOARD

© Copyright KALRAY SA. All rights reserved. Kalray, the KALRAY logo, MPPA, MPPA MANYCORE, MPPA DEVELOPER and MPPA ACCESSCORE are registered trademarks of KALRAY.



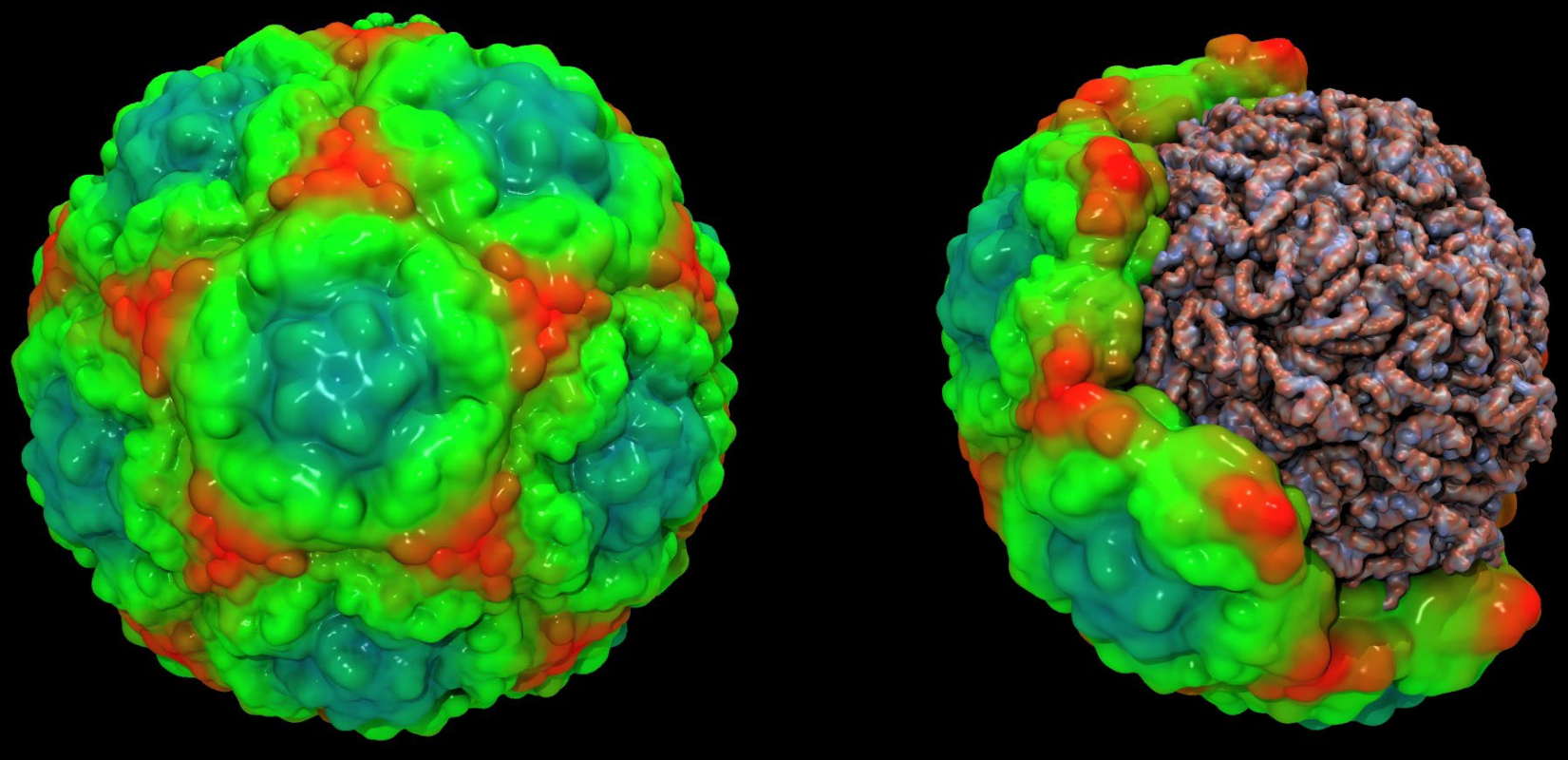
/en_couverture

HPC ET MEDECINE : DES AVANCEES DECISIVES

Recherche, pratiques de pointe, approches prédictives… la médecine connaît aujourd'hui une véritable révolution – le mot n'est pas trop fort – liée à la démocratisation du calcul intensif et des technologies Big Data. Les progrès sont exponentiels, les perspectives très prometteuses…

ALAIN CLAPAUD / FREDERIC MILLIOT

Conférence après conférence, les prix qui couronnent les travaux scientifiques majeurs font la part de plus en plus belle aux applications médicales du calcul intensif. Ce n'est pas un hasard. Les biosciences, l'e-Santé et le drug design sont parmi les premières à bénéficier des avancées concrètes du HPC. Exemple type, choisi pour son éloignement géographique afin que personne ne soit vexé, celui du St Vincent's Institute australien. En simulant, avec IBM Research et le soutien de l'Université de Melbourne, le développement d'un virus au niveau atomique, les chercheurs sont parvenus à faire "vivre" un rhinovirus sur Avoca,



Modéliser une entité biologique à combattre - ici un rhinovirus - pour simuler son comportement permet de tester des principes actifs in silico, avant de lancer les batteries de tests classiques in vitro et in vivo.

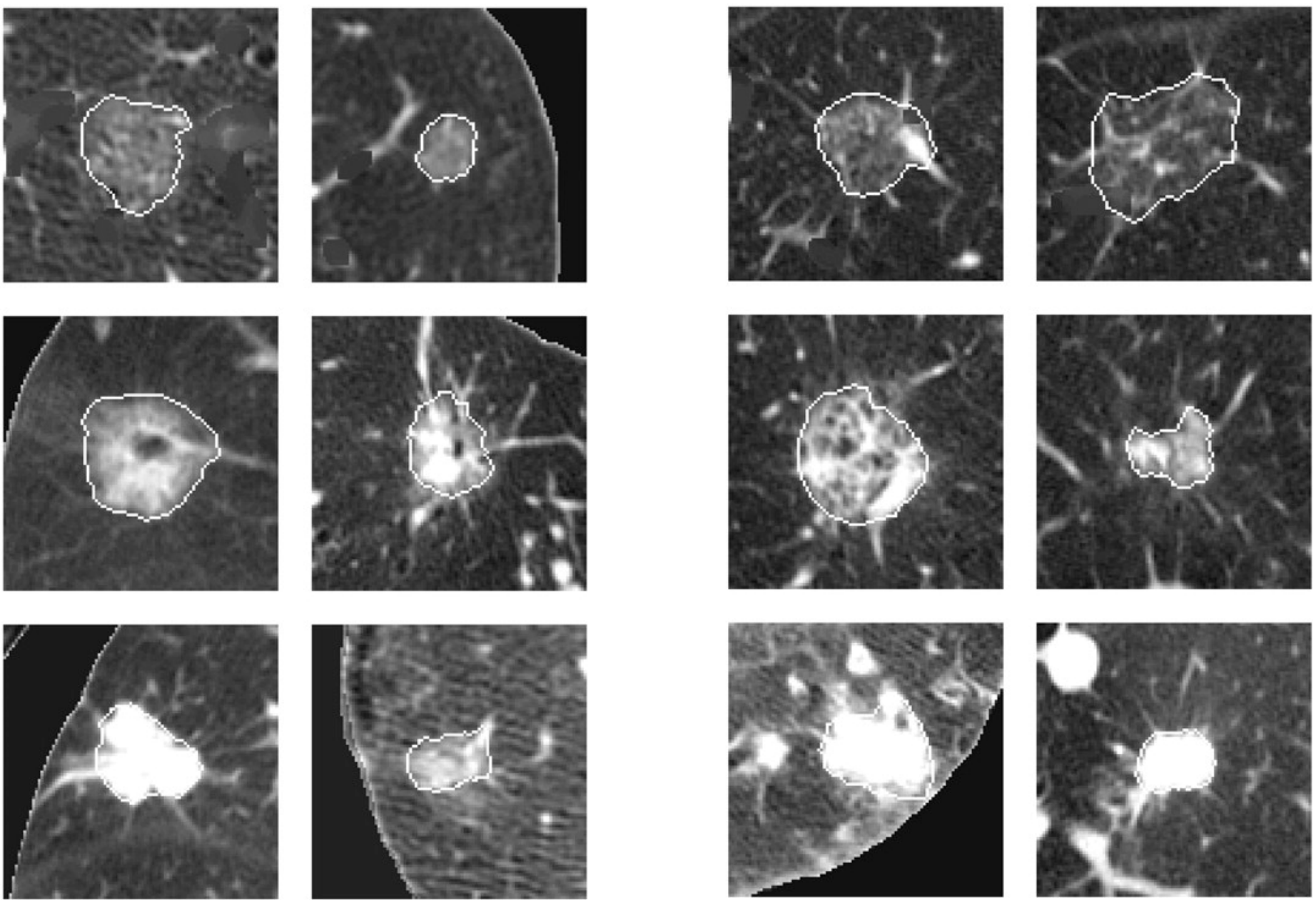
le plus puissant calculateur en production en Australie (Blue Gene/Q, 838 Tflops). Ce type de virus est responsable d'environ 40 % des rhumes. Il est par ailleurs impliqué dans 70 % des cas d'asthme. Le modèle élaboré simule non seulement les 3 millions d'atomes de l'enveloppe du rhinovirus dans son environnement aqueux, mais également son génome, à savoir un brin d'ARN comprenant 7 200 bases.

Ce faisant, l'équipe dirigée par Michael Parker cherche à valider virtuellement le comportement de nouvelles molécules curatives et préventives avant les tests in vitro et in vivo. L'utilisation de ce modèle est ainsi porteuse d'espoirs pour la santé de chacun d'entre nous, mais également synonyme d'économies qui, appliquées au seul domaine de la recherche, pourraient atteindre plusieurs dizaines de millions de dollars par an. A l'évidence, ces nouvelles méthodes ont tout pour séduire le secteur pharmaceutique. Le nombre de médicaments approuvés par les autorités de contrôle nationales, FDA en tête, décline en effet sensiblement depuis le milieu des années 90, précisément à cause de l'envolée des coûts de développement.

La fin du modèle réductionniste

L'approche "réductionniste" telle que la présentait le gériatre Philippe Guillet lors du forum Teratec 2013 est aujourd'hui devenue trop coûteuse. Fondée en grande partie sur le hasard - d'aucuns diront la sérendipité - elle consiste à sélectionner une molécule et à étudier ses effets sur des cultures de cellules, puis sur des animaux, puis si tout va bien à entamer une phase d'expérimentation sur l'homme. A chaque étape de "l'affinage", un certain nombre de variantes sont écartées en fonction des résultats cliniques observés et d'éventuels effets secondaires, de sorte que peu, au final, atteignent le stade final de la commercialisation.

La quête du *blockbuster*, capable de guérir tous les types de malades, est ainsi de plus en plus difficile à poursuivre pour les laboratoires. Elle devrait, progressivement, céder la place à approche plus systémique où les médicaments seront élaborés pour des groupes d'individus et personnalisés en fonction de leur génotype. Ce progrès, réel, n'est envisageable



Dans le cadre des applications de CAD (Computer Aided Diagnosis), les algorithmes d'analyse d'image permettent aujourd'hui de détecter certaines pathologies quasi automatiquement.

que grâce aux développement des méthodes et technologies de calcul - HPC pour la simulation multiéchelle et Big Data pour les analyses statistiques des données de test.

A terme, c'est une nouvelle vision de la médecine et de la relation médecin-malade qui se dessine, sur le principe des "4P" chers au Pr Luc Montagnier (prédiction, prévention, personnalisation, participation). Le patient ne se rend plus chez son médecin à l'apparition des symptômes, mais bien en amont, pour éviter de tomber malade. Détecter le plus tôt possible les probabilités pathologiques permet d'éviter leur apparition, ou si ce n'est pas le cas de les traiter avec des molécules sur mesure, évidemment mieux tolérées.

Les réseaux sociaux jouent les auxiliaires médicaux

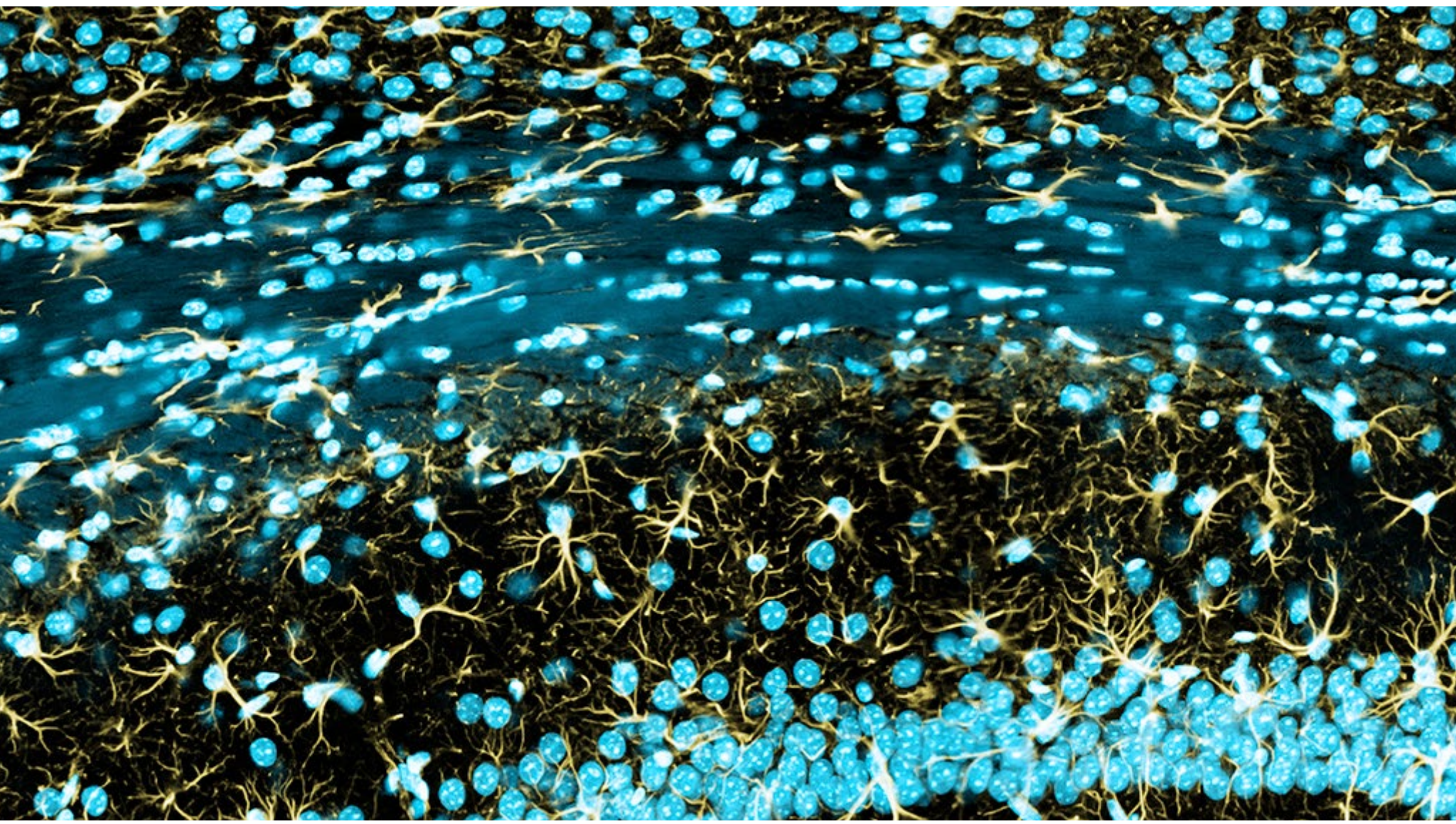
Quant au volet participatif, il fait appel au Web 2.0, y compris les médias sociaux au travers desquels l'ensemble de la population échange sur sa santé. Grâce à eux, le corps médical est à même de mieux impliquer les patients dans les prises de décision, mais aussi de détecter des tendances voire des épidémies avec une précision sans cesse accrue.

Cet aspect social peut paraître anecdotique et pourtant, outre qu'il améliore le niveau d'information des patients, c'est un outil de recherche en soi. En épidémiologie, ce sont les médecins et plus généralement les systèmes de santé qui remontent les informations liées aux maladies infectieuses et permettent ainsi de détecter les débuts d'épidémies. En France, le réseau Sentinelles, mis en place par l'Inserm et l'UPMC, collecte l'information auprès

d'environ 1 300 généralistes. La grippe, la rougeole et l'urétrite masculine sont ainsi suivies depuis 1984 - de nombreuses autres maladies infectieuses venant enrichir les bases de données du projet année après année.

Cette méthode, typiquement issue des pratiques Big Data, a clairement montré sa pertinence sur la grippe. Google l'a complétée d'un calcul prédictif qui permet au moteur de recherche de publier une estimation de l'avancée de l'épidémie

bilité de certaines pistes avant de déclencher le processus de recherche traditionnel. L'un des porte-drapeaux de cette approche en France est Michel Caffarel, directeur de recherche au CNRS, à la tête du laboratoire de chimie et physique quan-



Visualiser notre cerveau est un des objectifs du Human Brain Project. On voit ici une représentation de l'hippocampe à sa frontière corticale avec noyaux cellulaires (en bleu) et cellules gliales (en jaune). (Cliché Gabor Nyiri, IEM HAS)

Mais cet outil est aujourd'hui concurrencé par le Web. Depuis quelques années, Google a mis en place un système de suivi de la grippe et de la dengue à partir des recherches des internautes. Il se fonde uniquement sur les requêtes servies, partant du principe qu'un individu commence à rechercher des informations sur une maladie dès qu'il en ressent les premiers symptômes - ce que Google peut facilement repérer, localiser et comptabiliser. avec une précision que les spécialistes s'accordent à qualifier d'étonnante.

Simuler pour mieux soigner

Côté recherche, laboratoires et universitaires privilégient d'ores et déjà la voie des tests in silico en amont des traditionnels tests in vitro et in vivo. La raison ne vous aura pas échappé : simuler le comportement des molécules sur un calculateur permet d'explorer la via-

tiques (LCPQ). Les membres de son équipe sont à la recherche d'une alternative à la méthode DFT (Density Functional Theory / théorie fonctionnelle de la densité) pour résoudre l'équation de Schrödinger. A cette fin, ils partent d'une méthode probabiliste (cf. **notre dossier** à ce sujet) dans laquelle la position de chacun des électrons d'une molécule est établie à un instant T par tirage aléatoire. Cette approche permet de définir le comportement de la molécule

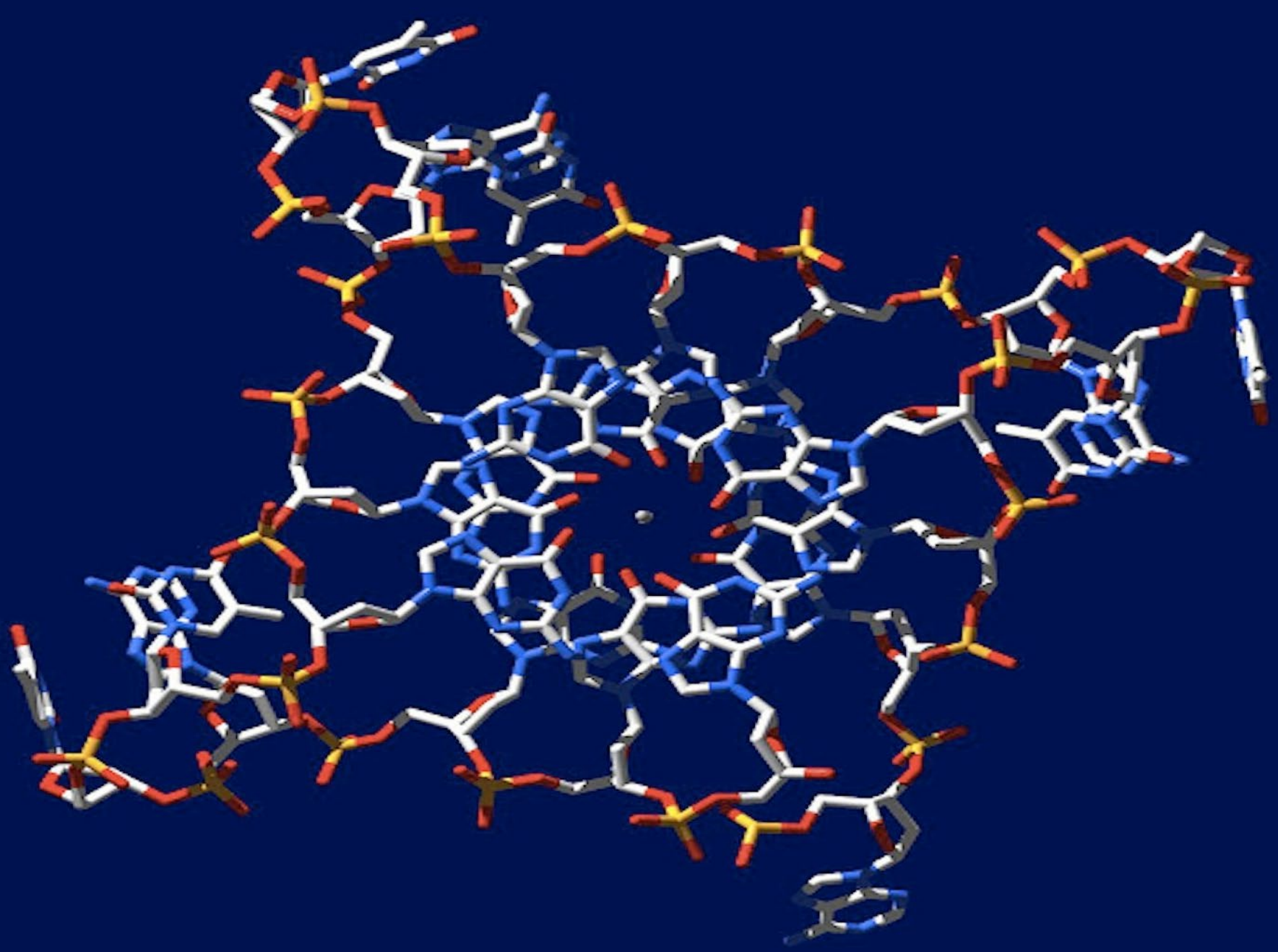
étudiée face à un ou plusieurs éléments externes. A ce titre, elle est particulièrement utile dans le cadre de la recherche contre la maladie d'Alzheimer.

Michel Caffarel a fait tourner ses modèles sur Curie, le calculateur du TGCC, pour étudier le comportement d'un peptide (la bêta-amyloïde) dans le cerveau. La molécule complète a ainsi été modélisée et le comportement des électrons de chacun de ses atomes calculé grâce à la méthode Monte-Carlo quantique, "facilement" parallélisable. En pratique, la manipulation a mobilisé l'essentiel des ressources de la machine pendant 2 nuits, soit une puissance consommée continue d'environ 960 Tflops. De quoi réaliser les 100 000 milliards de tirages aléatoires et obtenir la position des électrons à mesure qu'un atome de cuivre s'approchait du peptide. Une expérience remarquable au niveau mondial, qui montre combien la puissance de calcul détermine les progrès envisageables.

Sur ce plan, l'Europe a frappé fort avec le lancement de Human Brain Project, coordonné par l'EPFL de Lausanne et regroupant pas moins de 135 institutions, pour l'essentiel européennes. Nos lecteurs les plus assidus le savent, il s'agit rien moins que de simuler la structure et le fonctionnement de notre cerveau.

A la croisée de 3 domaines - les neurosciences, la médecine et le calcul neuroinformatique - cette simulation se place d'emblée à de multiples échelles. Moléculaire d'abord, avec protéines et chromosomes. Nerveuse ensuite, avec échanges neuronaux et synaptiques. Régionale enfin, pour affiner notre compréhension de la localisation fonctionnelle.

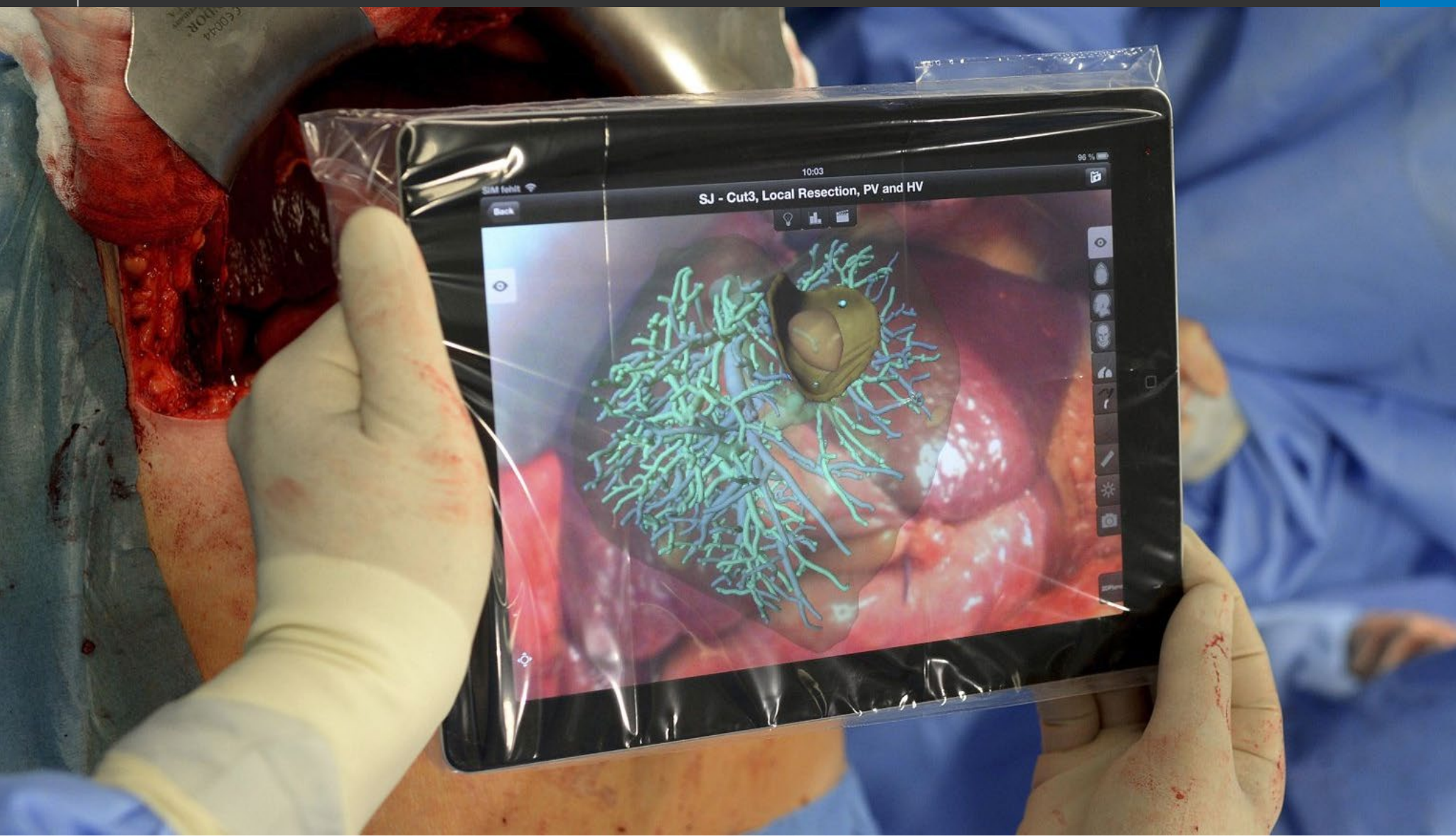
Offrir aux patients des traitements personnalisés basés sur le séquençage de leur ADN requiert des moyens de calcul conséquents. A terme, le développement de ces moyens devrait corréler assez finement avec l'augmentation de notre espérance de vie, d'autant que de récentes techniques permettent de modifier l'ADN à des endroits très précis.



Les fichiers issus de l'imagerie médicale représentent des masses de données colossales à stocker mais aussi une mine d'or en termes d'analyses Big Data.

Ces trois étapes devraient permettre d'envisager, dans un quatrième temps, une simulation globale. Le projet est pour cela séquencé en quatre étapes, à 4, 6, 7 et 10 ans, avec comme toile de fond une puissance de calcul qui va aller crescendo. Pour mémoire, les précédents projets du même ordre, dont **SpiNNaker**, se limitaient à la simulation d'environ 1 % des cent milliards de neurones et de leurs 10^{13} milliards de connexions.

Pour passer à l'échelle sur cette recherche, l'Europe s'est donné les moyens financiers adéquats, avec un budget en allocation directe supérieur au milliard d'euros. En phase de démarrage, HBP s'appuie sur JUQUEEN, la machine du centre de calcul de Jülich en Allemagne, un IBM BlueGene/Q de 5,9 Pflops, 450 To de mémoire centrale et 8 Po de stockage. Les plus gros calculateurs institutionnels européens sont mobilisés en soutien, dont ceux du CSCS, du BSC et du CINECA (sans oublier la plateforme NeuroSpin du CEA). Au total, une cinquantaine de Pétaflops et autant de Pétaoctets sont disponibles pour le *ramp up* du projet, sachant qu'au terme des dix ans de recherche, c'est bien d'un Exaflops que le cerveau simulé aura besoin pour tenter de rivaliser (même timidement) avec son modèle biologique. Outre le simulateur lui-même, le projet constitue une base de données de l'ensemble



Une opération chirurgicale effectuée en réalité augmentée, sur un iPad, à l'hôpital de Brème.

des publications scientifiques. Que ce cerveau synthétique s'éveille ou non à la conscience, HBP devrait donc permettre au Vieux Continent de progresser de manière significative dans le domaine de la simulation du vivant, d'améliorer sensiblement sa connaissance de notre cerveau et, accessoirement, de donner naissance à l'une des toute premières machine exaflopiques au monde.

Le Big Data, clé des traitements personnalisés

Pour beaucoup de scientifiques, la puissance de calcul est d'abord utilisée dans le but d'obtenir plus rapidement des données statistiques exploitables. Le but est ici d'offrir à chaque patient non plus un traitement type correspondant à sa pathologie mais une cure personnalisée et adaptée aux paramètres de son métabolisme et au stade de son affection. L'analogie peut sembler discutable, mais on peut techniquement comparer cette approche à celle d'un site Web marchand qui analyse le profil de ses clients en fonction de leurs achats passés et en déduit les produits les plus susceptibles de générer de nouvelles dépenses. De façon un peu similaire, un médecin devrait à terme disposer d'algorithmes analytiques sur les données de son patient pour décider des traitements et des doses les plus appropriés.

Les thérapies géniques illustrent de façon encore plus évidente l'avènement très prochain de nouvelles pratiques médicales. Depuis quelques années, on sait que certaines maladies comme l'hémophilie ou le diabète sont directement liées au patrimoine génétique des patients. Aujourd'hui, grâce aux **Genome Wide Association Studies**, on a déjà pu identifier 2 000 associations de gènes responsables de plus de 300 pathologies. L'intérêt de la méthode, clairement, est d'utiliser une "simple" analyse d'ADN pour administrer au patient un traitement capable de prévenir les maladies pour lesquelles il est programmé, avant même qu'elles ne se déclarent.

Très récemment, des chercheurs de l'université du Wisconsin et de l'hôpital Xavier Arnozan de Pessac en Gironde ont appliqué cette démarche pour guérir des chiens de l'hémophilie, maladie jusqu'ici considérée comme incurable. Les résultats sont probants mais la méthode a un coût - c'est même là son principal problème. Le séquençage d'un ADN humain, qui

compte entre 20 000 et 25 000 gènes, revient aujourd'hui à environ 5 000 euros. Les espoirs d'arriver rapidement à moins de 1 000 euros - un niveau sociétalement supportable - sont cependant réels. Rappelons qu'en 2007, il fallait compter 350 000 euros pour un séquençage complet… Il n'en reste pas moins qu'un génome numérisé représente 100 Go de données et qu'une généralisation du dépistage sur des populations de millions d'individus demeure un enjeu informatique et économique colossal.

Voilà pourquoi généticiens et bioinformaticiens se sont rapidement approprié la plateforme emblématique du Big Data - le framework Hadoop - pour leurs propres besoins de stockage et de traitement.

Dès 2009, Michael Schatz, de l'université du Maryland, proposait **CloudBurst**, une méthode de séquençage basée sur Map-Reduce. Selon son concepteur, la solution est 30 fois plus rapide que l'algorithme RMAP sur un cœur et, une fois déployée sur 96 cœurs sur Amazon EC2 par exemple, le temps de traitement passe de plusieurs heures à quelques minutes… Michael Schatz n'est pas le seul à avoir exploité Map-Reduce dans ce domaine : des alternatives telles que **Crossbow** ou **Hadoop-BAM** ont permis de très nettement simplifier l'implantation de ce type d'outil dans les laboratoires, et donc de démocratiser une pratique longtemps considérée comme académique.

Le M2M, principal fournisseur de données médicales

L'utilisation des technologies M2M (communications Machine to Machine) est elle aussi en plein boom. On la retrouve dans de multiples applications émergentes basées sur l'Internet des objets. Le M2M permet par exemple de suivre des patients de manière quasi continue entre les visites chez le médecin. Son émergence va amener les applications de télésanté à se multiplier, dans la mouvance du *Quantified Self* popularisée, entre autres, par Withings. Créée en 2008, la startup française a commencé timidement en développant une balance électronique accessible par smartphone et connectée à un site central où les utilisateurs peuvent suivre l'évolution de leur poids au jour le jour. Depuis, Withings a diversifié son offre vers le paramédical, avec notamment aujourd'hui un tensiomètre et demain des capteurs automatiques plus sophistiqués.

Toujours en France, Orange Business Services travaille avec le fournisseur d'équipements médicaux Weinmann dans le domaine du suivi de l'apnée du sommeil. Une solution, baptisée **homecareONLINE**, connecte le dispositif porté chaque nuit par le patient à une plateforme cloud qui stocke les données recueillies par les différents capteurs. La plateforme met ensuite les informations à disposition des personnels médicaux, via un portail à accès sécurisé.

Autre initiative de l'opérateur français, **M-Health Grand Tour**. L'opération avait pour but d'enregistrer une batterie de données sur les 150 participants d'une course cycliste organisée par la Fédération internationale de diabétologie (FID) et la GSM Association. Le quart des participants à cette épreuve était atteint de diabète de type 1 et faisait l'objet d'un télé-suivi en temps réel tout au long de la course via différents capteurs (fréquence cardiaque, pression artérielle, souffle, glycémie) ; certains participants avaient même été équipés d'une pompe à insuline. Les informations étaient transmises par un petit boîtier M2M intégrant une carte Sim, et consultables via une application mobile permettant aux professionnels de santé, aux chercheurs et aux coureurs de visualiser les données et les courbes historiques personnelles ou globales. Ces illustrations sont encore assez modestes, mais elles incluent l'essentiel des ressources technologiques permettant d'aller bien plus loin.

Existe-t-il encore une communauté au monde qui ne cherche à limiter l'explosion de ses dépenses de santé et à combler ses "déserts médicaux" ? Non, probablement, et c'est pourquoi le secteur de l'e-Santé est à la fois l'un des plus actifs et des plus innovants de la sphère IT. En France, **Cardiauvergne** en est l'exemple type. Soutenu par l'ARS d'Auvergne, le projet vise à exploiter une application de suivi des patients souffrant d'insuffisance cardiaque.

Connectée à une simple balance, l'application collecte chaque jour le poids et la fréquence cardiaque des patients, et les remonte directement sur la plateforme mise en place par Almerys (une filiale d'Orange). Plateforme à laquelle les laboratoires d'analyse médicale et les pharmacies de la région ont également accès, afin de compléter ces données. Au final, un dossier partagé centralise l'ensemble de ces informations, que médecins et personnels de soin peuvent consulter à tout instant.

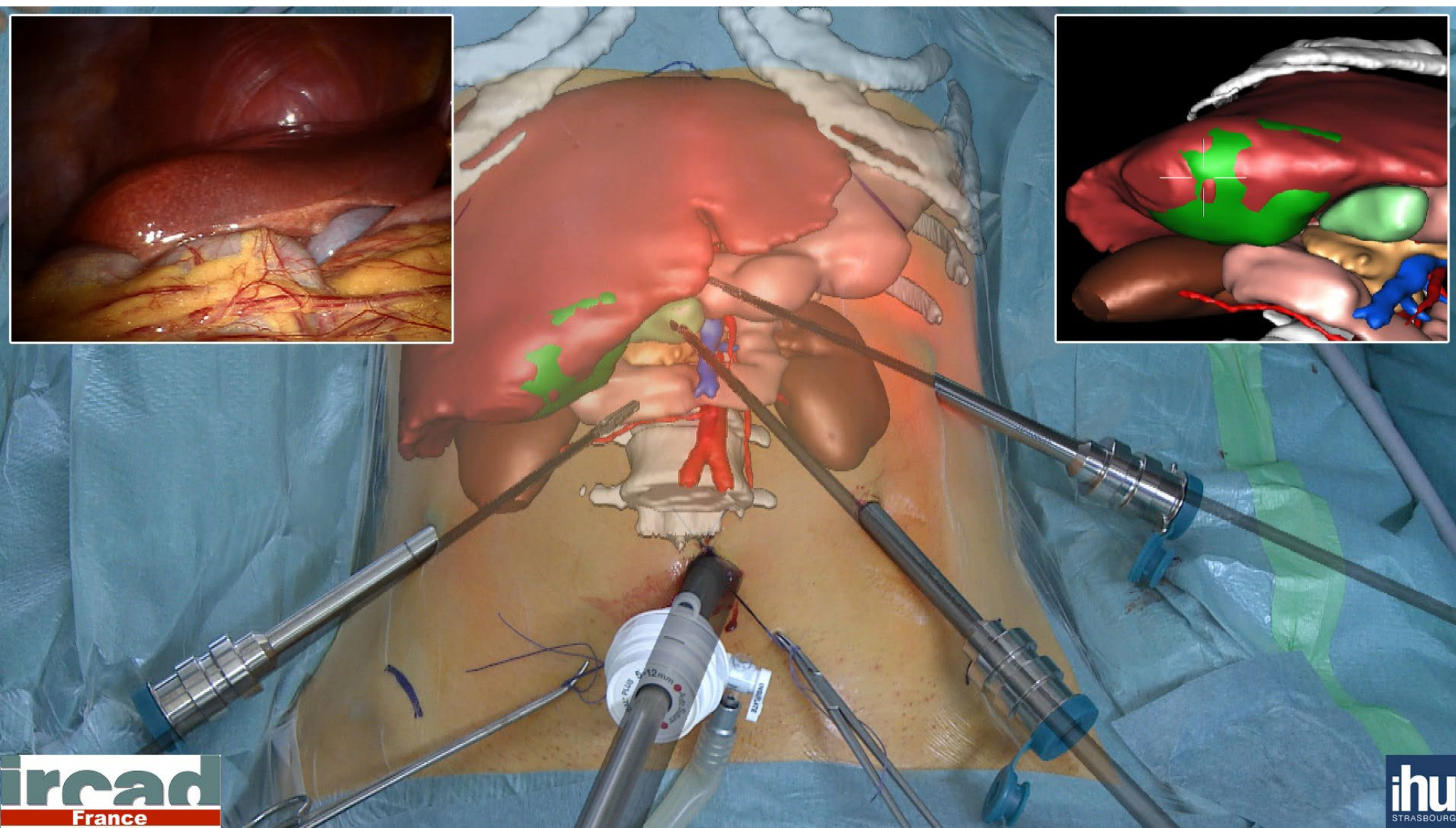
Le système est en outre capable de générer lui-même des alertes en cas de variations trop rapides de poids, de franchissements de fréquences cardiaques ou de résultats anormaux dans les examens biologiques. Lancé en septembre 2011, le programme pilote a permis de suivre 315 patients à ce jour. Quant au système d'alerte automatique, il a permis de générer 597 mises en garde sur une année, conduisant à 30 réhospitalisations. A terme, le projet devrait permettre de suivre 2 000 patients sur 3 ans et réduire les hospitalisations inutiles de 40 %, soit une économie moyenne effective de 7 000 euros par an et par patient.

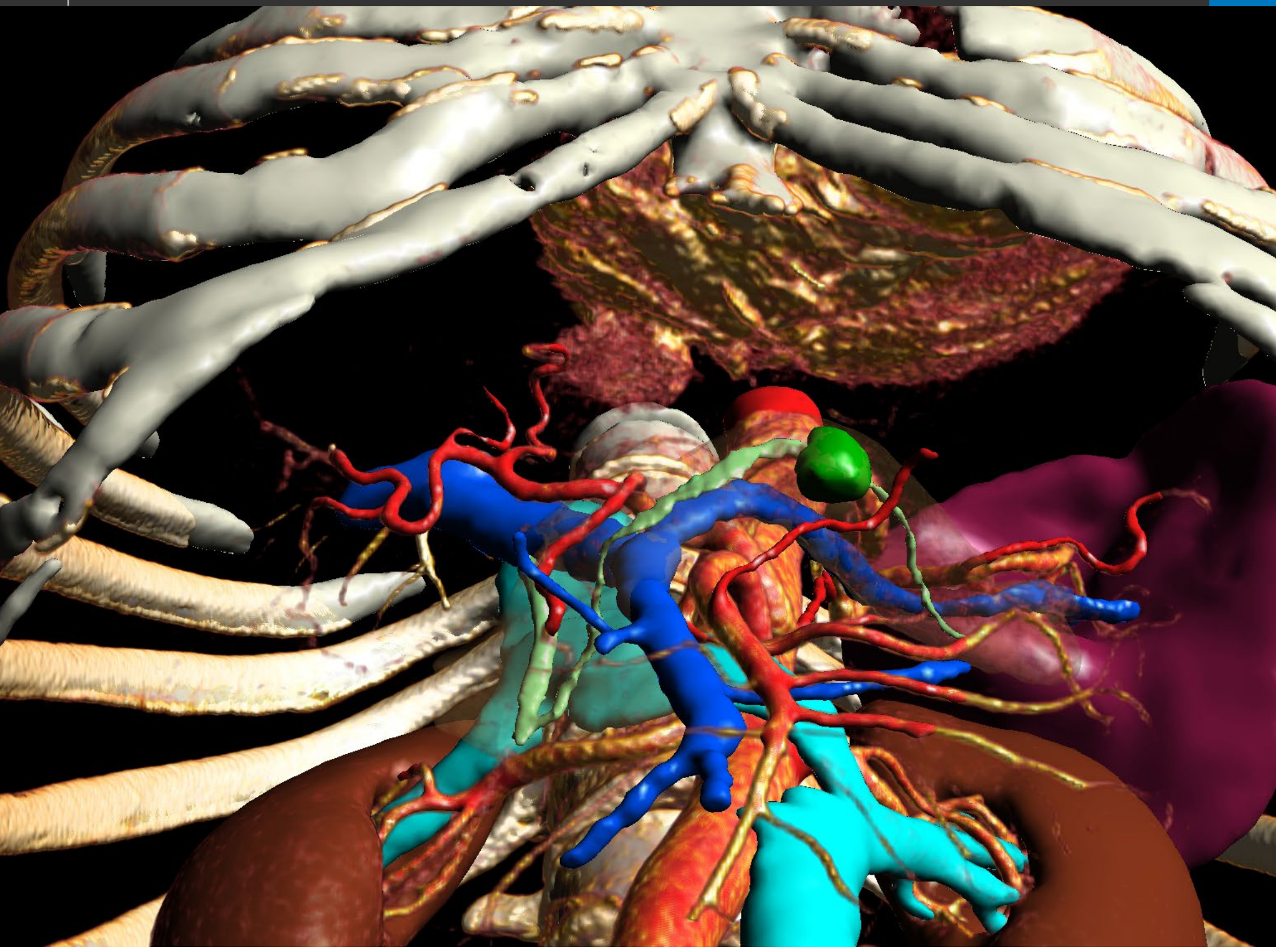
Cardiauvergne n'est qu'un exemple d'e-Santé parmi beaucoup d'autres. Dans le domaine privé, des sociétés américaines comme AirStrip, SHL Telemedecine ou Preventice proposent des patchs de poitrine à connecter à un smartphone qui transmet régulièrement l'électrocardiogramme des patients et déclenche des alertes en cas d'anomalie. Ce type de solution devrait très bientôt apparaître en Europe, et s'étendre à beaucoup d'autres usages que la cardiologie.

L'imagerie en totale mutation

S'il est un domaine médical dans lequel le Big Data s'est imposé d'emblée, c'est celui de l'imagerie médicale. Robots IRM, Pet Scans, les praticiens disposent de moyens d'imagerie de plus en plus sophistiqués et l'arrivée de la 3D n'a fait qu'ajouter à l'accroissement des volumétries à traiter. Pour y faire face, les professionnels s'organisent, à l'image du projet Région sans Film, une plateforme cloud

Basée sur la simulation, la réalité virtuelle permet aux chirurgiens de se former et de répéter des gestes délicats avant une opération complexe.

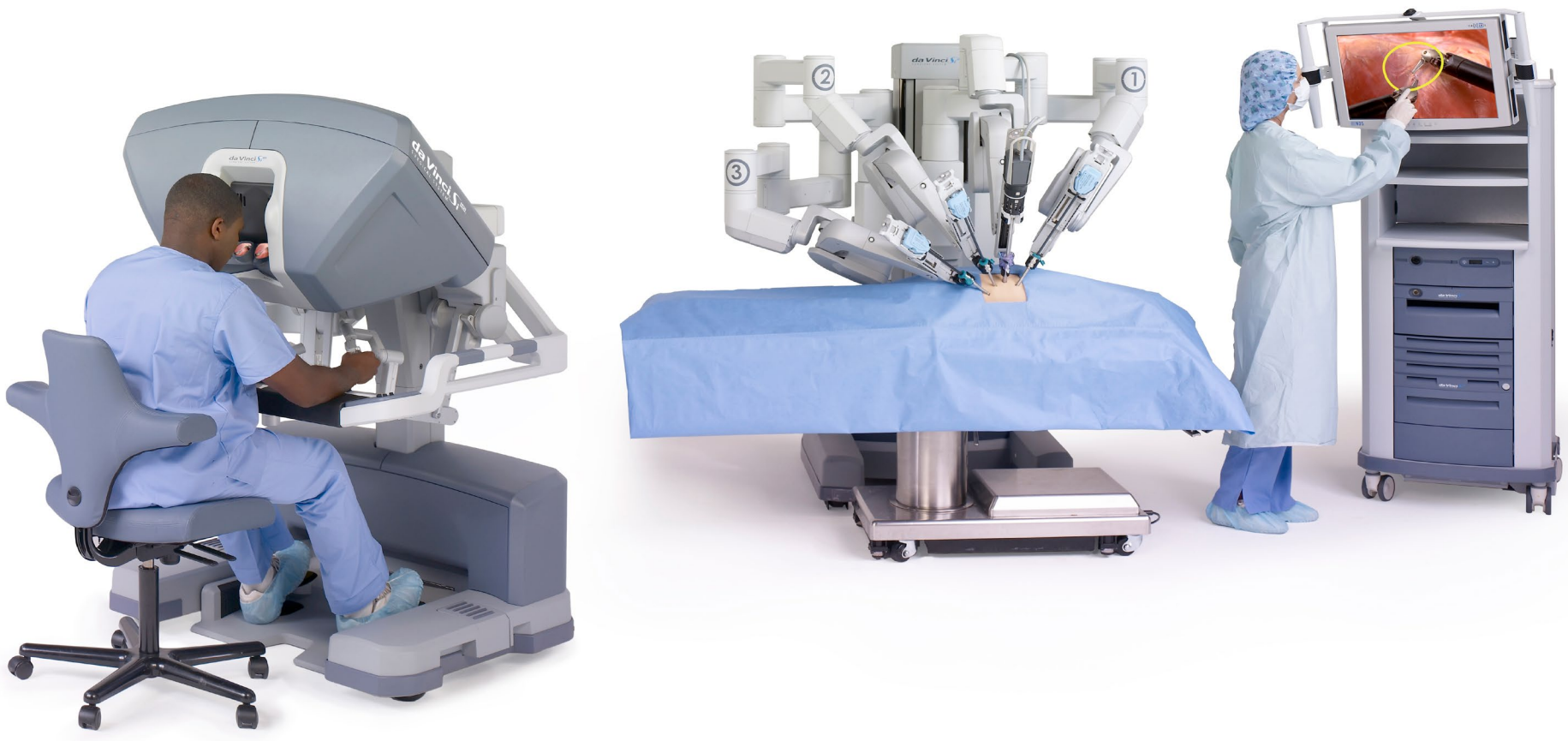




A partir de l'imagerie médicale, l'IRCAD génère d'époustouflants modèles 3D des organes de patients anonymes. Ces modèles sont exploitables ensuite par le chirurgien avant une intervention.

d'Imagerie médicale partagée (IMP) à disposition des établissements de santé et des médecins de ville. Son objectif : mutualiser les coûts liés au stockage et l'archivage sur le long terme. Son ambition : stocker jusqu'à 700 Pétaoctets dans des fichiers devant être conservés 15 ans !

Outre le volet partage et stockage de données graphiques, de multiples travaux sont menés depuis les années 80 dans le *Computer Aided Diagnosis*. L'idée est ici que la machine peut utilement seconder le praticien, en précision et en vitesse, dans la détection de nodules potentiellement cancéreux sur une radio ou de zones de métastases sur une scintillographie. A côté des algorithmes d'amélioration d'image et de reconnaissance de forme que l'on pourrait rapprocher de ceux utilisés pour l'analyse de clichés satellites, les chercheurs développent des fonctions très spécifiques. Exemple type : l'identification de pathologies cardiaques à partir d'une succession d'images IRM permettant de recréer la cinématique d'un cœur en train de battre. L'analyse des images tire très directement profit des techniques de programmation des FPGA et GPU mises au point sur les calculateurs académiques. Spécialisé dans les systèmes de dépistage des cancers du sein, l'américain Hologic, par exemple, exploite CUDA pour permettre aux radiologues d'obtenir en temps réel des clichés issus d'appareils d'imagerie médicale. Par rapport à un traitement CPU, le speedup atteint jusqu'à 250X. De plus, l'utilisation de GPU permet la miniaturisation des plateformes, ce qui les rend à la fois portables et compatibles avec une utilisation en salle d'opération.



Le robot chirurgical da Vinci, pionner de la chirurgie mini-invasive, offre aux praticiens une visibilité, une précision et une liberté de mouvements sans égale. Ce type de robot est appelé à se démocratiser, puis à se généraliser.

Les travaux de David Fuentes, du Centre de traitement du cancer MD Anderson de l'université du Texas, illustrent parfaitement cette tendance. En 2012, le Dr Fuentes a développé un code permettant de calculer l'échauffement des tissus cérébraux lors d'un traitement laser. Comme de multiples paramètres entrent en ligne de compte dans ce calcul (propriétés des matériaux, conductivité thermique, propriétés optiques, circulation sanguine…), ce code était suffisamment complexe pour mobiliser le calculateur Ranger du TACC. Un an après, des chercheurs de Rice sont parvenus à le porter en OpenCL pour le rendre exécutable sur un simple accélérateur GPU du marché. Là où l'obtention de résultats nécessitait une douzaine d'heures de calcul sur environ 12 cœurs CPU, 10 à 12 secondes y suffisent désormais.

Autre domaine extrêmement prometteur en termes de résultats cliniques, la modélisation 3D à partir d'images 2D. De nombreux laboratoires dans le monde travaillent sur le sujet, et notamment l'**IRCAD** (Institut de recherche contre les cancers de l'appareil digestif) du CHU de Strasbourg, créé par le Pr Jacques Marescaux.

Avec ses antennes à Taïwan et au Brésil, l'institut forme chaque année 3 500 praticiens à la chirurgie mini-invasive. Son centre d'excellence, animé par une équipe de R&D de pointe, travaille sur des techniques de réalité virtuelle et de réalité augmentée pour former les médecins et les assister lors des opérations.

Parmi ses réalisations les plus emblématiques, il faut nécessairement citer **Visible Patient**, un ensemble de logiciels et de modélisations du corps humain incluant une bibliothèque librement accessible de 2 000 modèles 3D correspondant à des cas cliniques anonymes. Médecins et chirurgiens peuvent y télécharger des modèles, les consulter avec l'application de visualisation développée en interne et étudier en trois dimensions les pathologies auxquels ils vont avoir affaire. Ils peuvent également y transférer au format standard DICOM (Digital Imaging and Communications in Medicine) les images IRM ou de tomographie numérique de leurs patients, et les comparer aux modèles 3D correspondants. Le praticien dispose ainsi d'une panoplie d'outils très performants pour simuler l'opération dans une situation aussi proche que possible de la réalité et pour la répéter autant de fois que nécessaire, jusqu'à maîtriser parfaitement le geste à accomplir.

Le chirurgien est un robot

C'est un fait, la 3D est désormais entrée dans la salle d'opération. Microsoft Research a ainsi mené un projet visant à permettre au chirurgien de manipuler des modèles 3D et de consulter les éléments du dossier du patient sur un écran en salle d'opération, directement via l'interface gestuelle d'une Kinect. Le praticien peut ainsi rester en place. Il n'a plus à quitter le bloc pour aller vers un poste informatique, ni à manipuler la souris pour consulter les données dont il a besoin.

En août dernier, des chirurgiens de l'hôpital de Brême sont allés beaucoup plus loin. Assistés par les chercheurs du laboratoire MEVIS de l'Institut Fraunhofer et de l'hôpital universitaire de Yokohama, ils ont opéré un patient en s'aidant d'une tablette iPad et d'une application de réalité augmentée ! La tablette affichait le modèle 3D du système veineux du foie du malade en direct. Les veines, invisibles par les chirurgiens, venaient se superposer à l'image vidéo du foie délivrée par l'iPad, pour leur permettre d'éviter autant que possible d'endommager des vaisseaux et de prévenir tout risque d'hémorragie grave.

Ces aides visuelles s'ajoutent à l'arsenal robotique que l'on trouve également de plus en plus souvent dans les hôpitaux pilotes. Jusqu'à aujourd'hui, les robots d'intervention chirurgicale tels que da Vinci fonctionnent sur le modèle de la télécommande : les instruments

Le robot Rosa du français MedTech assiste le neurochirurgien au cours de l'opération, mais ne le remplace pas.

miniaturisé reproduisent les gestes que le chirurgien exécute sur une console, le robot fournissant en plus une vue 3D du champ opératoire. En d'autres termes, ce n'est pas le robot qui opère lui-même le patient. Il assiste le praticien par voie d'imagerie pour en quelque sorte guider sa main.

Un projet de la NSF américaine lancé en 2012 a pour ambition de dépasser ce simple assistanat. Doté d'un budget de 50 millions de dollars sur 4 ans, mené conjointement par des chercheurs des universités de Stanford, Berkeley, Johns Hopkins, Washington et Santa Cruz, il vise à développer des robots collaboratifs qui pourraient assister les chirurgiens de façon beaucoup plus active.

Si les implications, notamment au niveau légal, sont loin d'être résolues, des actes simples peuvent déjà être réalisés de manière autonome par une machine. Dès 2009, les bioingénieurs de Duke ont démontré l'efficacité d'un robot capable de localiser par lui-même des éclats d'obus dans le corps d'une victime et de guider une aiguille pour les en extraire. D'abord destinée à une utilisation sur les théâtres d'opérations armées, cette solution devrait également trouver des applications à l'hôpital, notam-

ment en cancérologie pour le placement de sources radioactives au plus près des cellules malignes.

Diagnostic :
le dernier mot à l'ordinateur

Chacun connaît Watson, le calculateur d'IBM vainqueur du célèbre jeu télévisé américain Jeopardy. Big Blue, en association avec le centre Sloan-Kettering de New-York et l'entreprise américaine Wellpoint, ambitionne de le proposer aux hôpitaux non pour la recherche ou la préparation d'actes médicaux, mais comme outil de diagnostic.

On le sait, Watson est constitué d'un mix de technologies d'indexation de données (IBM Content Analytics), de calcul prédictif issus de l'éditeur SPSS, et de Big Data avec Big Insight, la distribution Hadoop de la marque. IBM a donc fait avaler à son calculateur 600 000 actes médicaux, 2 millions de pages d'articles scientifiques relatifs au cancer, mais aussi et surtout 1,5 million de dossiers médicaux de patients. Bien plus que ce qu'un médecin lira jamais. Résultat, cette masse d'information restructurée permet à la machine, lorsqu'on lui présente le dossier d'un nouveau patient, d'obtenir des diagnostics plus fiables que ceux d'un praticien humain.

Selon WellPoint, Watson livre un diagnostic correct dans 90 % des cas pour un cancer de la langue - une performance bien supérieure à celle des praticiens spécialisés dont le taux de diagnostic correct dépasse rarement 50 % pour ce type de pathologie. Alors, à quand un calculateur Watson dans chaque service d'oncologie ?

Il va falloir attendre un peu car le coût d'acquisition et d'exploitation de cette machine - dont on assure chez IBM qu'elle est tout ce qu'il y a de plus standard (il s'agit en fait de serveurs IBM P Series 750) – la met encore hors de portée des hôpitaux les mieux gérés. IBM propose donc aux équipes d'accéder à ce service de diagnostic via le cloud, en mode locatif.

Sur le papier l'offre est financièrement intéressante, d'autant que Watson présente la particularité de pouvoir être interrogé en langage naturel. On peut donc imaginer les cancérologues interagir avec le calculateur sans trop de peine. Et si IBM a commencé par le cancer, rien n'empêche de faire ingurgiter à la machine d'autres types de dossiers médicaux. Le hic c'est que, contrairement à l'imagerie ou la simulation, la technologie mise en œuvre par Watson est fortement dépendante de la langue en laquelle sont rédigés les documents indexés, et qu'il n'a encore rien démontré en français, en allemand ou en espagnol. Qui aurait un jour imaginé que la médecine à deux vitesses serait fonction de notre langue maternelle... ■

NVIDIA Quadro K6000,

le GPU le plus rapide et le plus avancé jamais conçu

Plus d'informations sur produits, visitez www.pny.eu/quadro
et le configurateur PNY http://pny.quadrok-selector.com/

Suivez-nous: @PNYProFR

PNY®

© 2013 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, NVIDIA Quadro, Tesla, and CUDA are trademarks and/or registered trademarks of NVIDIA Corporation. All company and product names are trademarks or registered trademarks of the respective owners with which they are associated.

NVIDIA.

/verbatim

JEAN-FRANÇOIS LAVIGNON

DIRECTEUR STRATEGIE ET COOPERATIONS - BULL
CHAIRMAN - ETP4HPC

PROPOS RECUEILLIS PAR STEPHANE BIHAN

A l'ordre du jour :

- L'ETP4HPC entre "dans le dur"
- Les enjeux de l'indépendance technologique européenne
- Extreme Factory récompensée à SC'13
- Des API pour parvenir à l'Exascale

Monsieur Lavignon, en tant que Président d'ETP4HPC (European Technology Platform for HPC), pouvez-vous nous en rappeler brièvement l'organisation, la mission et les objectifs ?

J.-F. Lavignon : L'ETP4HPC a pour mission de discuter avec la Commission sur les travaux de recherche à mener dans un domaine donné, en l'occurrence celui du calcul haute performance. Cette entité réunit l'ensemble des acteurs concernés, définit les priorités qui lui semblent importantes et communique avec la Commission pour monter des programmes visant au dévelop-

> ***On ne peut plus se contenter d'utiliser des machines ou des logiciels produits par les autres. L'Europe a besoin d'un tissu de compétences à tous les niveaux : technologies, utilisateurs et services.***

pement du secteur. En termes d'organisation, l'ETP est une association de droit hollandais qui comprend aujourd'hui une cinquantaine de membres. Elle est dirigée par un *board* au sein duquel siègent des représentants d'Allinea, ARM, BSC, Bull, CAPS Entreprise, CEA, CINECA, Eurotech, Fraunhofer, IBM, Intel, Jülich, ParTec, STMicroelectronics et Xyratex.

Et combien de membres visez-vous à terme ?

L'ensemble des *stakeholders* peut représenter une centaine d'organisations réellement impliquées dans la recherche en calcul intensif en Europe.

Nous souhaitons également tisser des liens, pas obligatoirement de type *membership*, avec des utilisateurs de toute taille, de façon à ce qu'ils soient bien informés de ce que nous faisons et qu'ils nous retournent leurs préoccupations. C'est une façon pour nous de mieux nous connecter avec les consommateurs de nos recherches.

L'ETP4HPC a défini un SRA (Strategic Research Agenda) qui trace les grandes lignes d'un programme destiné à renforcer la position européenne au niveau mondial. Quels en sont les axes principaux ?

La première étape a logiquement consisté à définir une vision de ce que l'on souhaitait accomplir. Cette vision est claire : le constat est que le domaine du calcul haute performance est véritablement stratégique, tant pour la compétitivité économique de l'Europe que pour son excellence scientifique. Aujourd'hui, pour être en pointe dans le HPC, il faut être en pointe dans ses technologies. On ne peut plus se contenter d'utiliser des machines ou des logiciels produits par les autres et clamer ensuite qu'on sera en avance au niveau utilisation. L'Europe a besoin d'un tissu de compétences à tous les niveaux, c'est à dire *grosso-modo* technologies, utilisateurs et services. Pour ce qui concerne plus particulièrement les technologies, les opportunités sont là : on n'avancera pas éternellement sur la seule base du renouvellement de l'existant. D'importantes évolutions sont à prévoir, qui devraient permettre aux acteurs européens d'acquérir de nouvelles parts de marché.

Pour atteindre notre ambition de construire un écosystème HPC solide et qui puisse à terme autofinancer son développement, le SRA a identifié quatre dimensions à développer. Il nous a semblé important de viser les technologies permettant d'atteindre l'Exascale, mais pas uniquement. Sur un plan économique, l'Exascale est un marché étroit – en tous cas par rapport à nos ambitions d'autofinancement.

Cette dimension a bien évidemment un rôle moteur, mais elle n'est pas la seule. Le SRA indique d'ailleurs clairement que la vocation première des recherches à mener porte sur les technologies de base du HPC. A ces deux dimensions complémentaires qui bornent notre horizon s'en ajoute une

troisième : le développement de nouveaux modes d'usage. Nous pensons que le HPC va jouer un rôle de plus en plus critique dans un certain nombre de systèmes. C'est ce que l'on appelle le HPC *in the loop*, c'est-à-dire le contrôle de systèmes très complexes avec beaucoup de données à analyser et à optimiser en temps contraint. Quand on considère également tous les développements dans le domaine Big Data, on voit que les usages du calcul intensif sont en train de se diversifier. Comme ce sont de nouvelles pratiques, il est important d'investir prioritairement dans ces domaines en Europe, pour nous positionner au meilleur niveau mondial.

Le dernier axe vise à la création d'un environnement global favorisant la réalisation de nos objectifs, notamment en termes de formation et d'incitation à l'innovation et à la création d'entreprises. Si l'on veut réussir, cet effort est absolument indissociable de la recherche pure. Voilà, en résumé, les grandes lignes du SRA. Elles y sont déclinées en 6 thèmes techniques incluant pour chacun d'eux une séquence de jalons et des deadlines. L'ensemble est détaillé dans le document officiel téléchargeable sur notre site Web.

PRACE est en charge de fournir l'e-Infrastructure HPC en Europe. Comment se positionne ETP4HPC dans ce contexte ?

Nous sommes très complémentaires. L'Europe a besoin d'être forte dans les technologies mais également d'offrir à sa communauté scientifique les accès les plus performants aux moyens HPC dont elle dispose. PRACE permet cela. Elle offre aux chercheurs du Vieux Continent les meilleurs outils de simulation, ce qui nous permet de maintenir notre leadership mondial dans ce secteur. L'ETP est très favorable à cette dualité, qui par ailleurs favorise les retours des utilisateurs sur les technologies actuelles.

Il est d'ailleurs probable que les technologies que nous allons développer, sur la base de ces retours utilisateurs, enrichiront l'infrastructure de PRACE dans le futur. Cette notion d'apport mutuel est fondamentale. En partageant notre vision des évolutions à venir, nous espérons permettre par exemple aux scientifiques qui produisent leurs

> ***La Commission se dit prête à engager 700 M€ de financement sur 7 ans, et souhaite que les autres acteurs contribuent à même hauteur...***

propres codes de prendre les bonnes décisions en matière de redesign applicatif.

Dans le cadre du programme Horizon2020, l'Europe compte investir 150 M€ par an à travers la mise en place de cPPP (contractual Public Private Partnership). Que sont ces cPPP exactement et quel est le rôle de l'ETP4HPC dans leur mise en place ?

En fait, ce que nous préconisions dans le SRA, c'est que l'effort global de recherche européen sur le HPC soit d'au moins 150 M€ par an. La Commission a pour sa part estimé qu'un engagement de ce type sur le long terme nécessitait un cPPP. Aux termes de celui-ci, la Commission se dit prête à engager 700 M€ de financement sur sept ans, soit 100 M€ par an, et souhaite que les autres acteurs contribuent à hauteur équivalente. Ce qui reviendrait au total à 200 M€ par an pour les technologies HPC et les centres d'excellence chargés de développer les applicatifs. Nous pensons qu'il existe suffisamment d'acteurs prêts à investir dans ce secteur pour que ce niveau global de 200 M€ annuels puisse être atteint.

Pour revenir plus précisément sur les cPPP, ce sont des structures très légères où siègent des représentants de la Commission et des acteurs de l'écosystème. Le rôle opérationnel d'un cPPP est de permettre les échanges entre ces représentants ; c'est d'ailleurs dans ce cadre que l'ETP va alimenter le programme de la Commission par des propositions. Le SRA en constitue la première. Ensuite, la Commission met en place des programmes qu'elle gère dans Horizon2020 de la même façon qu'elle les gérait dans le FP7. C'est-à-dire qu'elle va publier des appels auxquels les consortiums répondront.

Y a-t-il des différences par rapport au FP7 dans les modalités de financement ?

Oui, la Commission a désormais tendance à vouloir financer à 100 % les projets qu'elle sélectionne. Ce qu'elle attend en contrepartie, c'est que d'autres actions soient directement prises en charge par le financement privé de façon à ce que, globalement, l'effort de recherche soit à la hauteur des 200 M€ envisagés par an. Désormais le contrat est signé entre l'ETP et la Commission et c'est cette dernière qui publie les programmes et qui organise la sélection, avec une recommandation de notre part.

A quelle échéance ces programmes seront-ils arrêtés ?

Un appel devrait être publié en décembre, auquel il faudra répondre en novembre 2014 avec des propositions. A partir du draft publié par la Commission Européenne et les Etats membres, nous avons aujourd'hui une vision assez précise du contenu de l'appel.

La Commission évoque également des centres d'excellence. Quelle sera leur mission concrète et comment vont-ils se matérialiser ?

Sur ces centres d'excellence, nous sommes moins avancés que sur le programme de R&D. Des choses peuvent encore évoluer mais l'idée directrice est de constituer des entités qui préparent les applications de demain et rassemblent les compétences dans un secteur donné. Là encore, la Commission souhaiterait travailler par appel à propositions. Ce sont des consortiums plus orientés recherche qui vont répondre en présentant des

projets pour développer, par exemple, des codes de combustion ou des applicatifs liés aux sciences de la vie. Dans un premier temps, on espère voir naître entre quatre et cinq centres, pour aboutir à une dizaine à terme. Chacun aura pour objectif, en plus du développement logiciel, d'animer le secteur en partenariat avec les utilisateurs. Il s'agit là d'actions de formation et de développement de services destinés à la diffusion des applicatifs et des bonnes pratiques HPC qu'elles impliquent.

La capacité à concevoir et à fondre les processeurs est un enjeu crucial pour le HPC et pour l'embarqué. Dans ce domaine, largement dominé par nos amis américains, comment l'Europe peut-elle gagner en indépendance et en compétitivité ?

Ces aspects sont des enjeux de dimension mondiale. L'Europe ne cherche pas à fermer son marché ou à développer des solutions régionales qui pourraient ne pas être au meilleur niveau par rapport à la compétition. L'idée est plutôt d'avoir suffisamment de poids pour être influent dans ce secteur. Si les processeurs sont fondus aux Etats-Unis mais qu'en Europe on a la vision de ce qui va se passer dans ce domaine très particulier, ce n'est pas très gênant. L'essentiel, c'est de pouvoir anticiper les évolutions. On peut être contributeur à un certain niveau même si toute la valeur ajoutée n'est pas en Europe.

L'ETP n'a d'ailleurs pas pour idéal l'indépendance totale de l'Europe sur toutes les briques technologiques. Ce serait assez utopique dans une économie mondialisée. Il nous faut juste suffisamment de compétences sur l'ensemble de la chaîne de valeur pour comprendre les implications de ses évolutions et pouvoir en tirer profit sur certains créneaux. En d'autres termes, pour développer une valeur ajoutée en Europe et être au meilleur niveau mondial. Par ailleurs, je pense personnellement qu'il y a des opportunités dans les processeurs en Europe, mais qu'il ne faudra pas considérer comme un échec le fait que la majorité des processeurs continue à être développée en dehors du continent.

Il n'y pas donc pas finalement de criticité à l'indépendance ?

Faire de l'indépendance pour de l'indépendance, ça n'a pas beaucoup de sens. Dans notre secteur, il faut vraiment concevoir l'effort de recherche avec une vision économique. Si le contexte ne permet pas la profitabilité, il n'est probablement pas nécessaire d'investir au delà de la compétence suffisante pour avoir la visibilité de là où on veut aller.

La compétitivité de l'Europe dans le HPC doit permettre aux utilisateurs d'être vraiment à la pointe. La meilleure façon d'y parvenir, c'est d'avoir le tissu et l'écosystème...

La France est un des quatre pays au monde, avec les Etats-Unis, la Chine et le Japon, à être active sinon indépendante sur l'ensemble de la chaîne de valeur du HPC. Voyez-vous cet état de fait perdurer et, plus globalement, a-t-il encore un sens à l'heure où la dimension régionale européenne prend une importance croissante ?

A nouveau, on ne cherche pas à être indépendant pour être indépendant. On cherche à avoir une compétence suffisante pour être maître de notre destin. Cela dit, au risque de me répéter, des arbitrages sont nécessaires en fonction de critères économiques. Si, à un moment donné, il est plus intéressant d'utiliser des briques qui viennent d'ailleurs, il faut le faire, en prenant soin de garder une bonne compréhension des évolutions technologiques. De cette façon, si les conditions économiques venaient à changer, s'il devenait indispensable de nous remettre à produire de façon autonome, nous serions en capacité de le faire, sans avoir besoin de personne. C'est un point vraiment fondamental.

Puissance de calcul et indépendance technologique sont des enjeux majeurs pour les grands blocs que sont notamment l'Europe, la Chine et les Etats-Unis. Mais pensez-vous que les efforts technologiques, donc financiers, nécessaires pour les prochains passages à l'échelle puissent se concevoir autrement qu'en coopération internationale ?

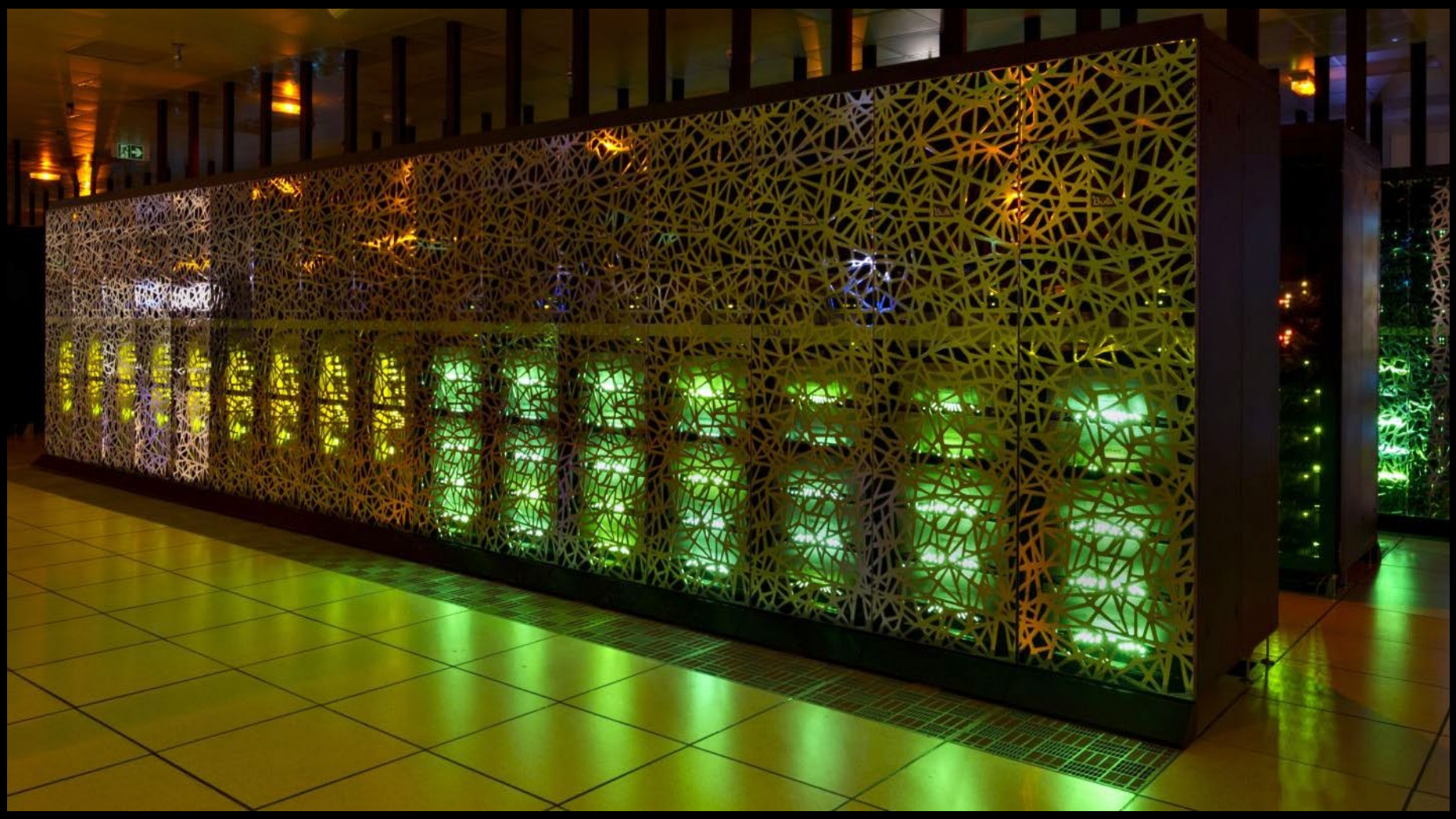
Ce qu'il faut bien comprendre c'est que la compétitivité de l'Europe dans le HPC doit permettre aux utilisateurs d'être vraiment à la pointe. S'ils attendent des solutions technologiques qui viennent de l'extérieur, ils ne peuvent pas prétendre être en avance par rapport à des concurrents situés dans les zones géographiques où ces technologies sont développées. Par exemple, pour les utilisateurs de codes automobiles, aéronautiques ou *oil & gas*, il est indispensable de pouvoir anticiper les évolutions. Et la meilleure façon d'y parvenir c'est d'avoir en Europe le tissu, l'écosystème nécessaire pour leur permettre d'avancer dans leurs usages. Ne pas avoir cette capacité en Europe entraînerait un retard sur l'ensemble de la chaîne, avec pour conséquence des effets sur la compétitivité globale de l'économie européenne.

Outre la présidence d'ETP4HPC, quelles sont vos fonctions chez Bull ?

Je suis en charge de la stratégie et des coopérations de recherche de la *Business Line* "Innovative Products". C'est cette entité qui conçoit les serveurs pour le calcul haute performance mais aussi ceux que Bull propose au monde de l'entreprise.

Pour information, nous avons aujourd'hui une activité soutenue sur des serveurs x86 relativement puissants qui sont très bien adaptés au traitement de données. Le Big Data est ainsi un créneau sur lequel Bull se développe actuellement.

La rédaction tenait à féliciter Bull Extreme Factory qui a reçu, pendant SuperComputing 2013, le prix de la meilleure technologie de visualisation 3D à distance pour la simulation numérique.



En quoi Extreme Factory se distingue-t-elle des offres concurrentes dans le domaine du HPC à la demande ?

L'originalité d'Extreme Factory, en dehors de ses technologies d'infrastructure, c'est l'intégration des offres de nombreux ISV : Altair, Ansys, CD-adapco, CST, ESI Group, Simulia/DS, The Mathworks… Dans ce domaine, Extreme Factory propose plusieurs types de services à travers un portail géré par Bull. D'abord, le portail permet aux utilisateurs d'accéder directement à des ressources de calcul à la demande. Ensuite, les ISV ont la possibilité de personnaliser le portail pour proposer leurs propres offres de calcul à distance. Enfin, une troisième déclinaison permet à un client d'installer ses propres codes et de se faire son cloud privé. Cette solution, baptisée Extreme Computing Studio, est installée *in situ*. Dans Extreme Factory, tout est intégré pour que les gens puissent travailler sur Radioss, Fluent, ANSYS Mechanical, Star-CCM+, STAR-CD, Virtual Performance Solution, etc. aussi simplement que si l'application était installée en interne. Nous avons pour cela tous les gestionnaires de licences nécessaires. Soit les utilisateurs ont déjà des jetons qu'ils peuvent transférer et utiliser à travers Extreme Factory, soit ils achètent des jetons directement sur Extreme Factory.

De même, en quoi les systèmes Bullx pour le HPC se distinguent-ils de leurs concurrents directs ?

Comme la plupart de nos concurrents, nous travaillions avec des CPU Intel et des accélérateurs Intel ou NVIDIA. En revanche, c'est dans l'intégration des nœuds de calculs que nous innovons. Nos solutions étant conçues spécifiquement pour le HPC, nous poussons cette intégration assez loin, en n'embarquant que ce qui est réellement nécessaire au calcul. Nous avons également des solutions de refroidissement performantes, soit à air, avec éventuellement une porte à eau pour éviter de climatiser la salle, soit en direct liquid cooling (DLC), avec de l'eau tiède, qui donne un excellent PUE. Globalement ces solutions présentent un TCO très intéressant pour le client, avec une haute fiabilité et une excellente maintenabilité.

Nous mettons également l'accent sur l'environnement logiciel. Notre palette d'outils de gestion intégrés offre une vision globale de tout le système, stockage y compris, et permet d'optimiser les charges de travail pour l'efficacité computationnelle et/ou énergétique. Nous avons

Les développeurs applicatifs vont devoir utiliser des API pour exprimer la sémantique de leurs codes et travailler indépendamment des modèles fins des machines...

aussi beaucoup travaillé sur la gestion des données. Bull est sans doute un des acteurs européen les plus avancés dans la maîtrise de Lustre. Ces spécificités nous permettent de proposer aussi bien des solutions de milieu de gamme très efficaces que de très grandes installations sur lesquelles nos clients ont des taux de disponibilité supérieurs à 95 %.

Bull est présent principalement en Europe. Sur la région Amériques, quelle stratégie de développement déployez-vous, notamment sur le plan technique ?

L'Europe est le marché principal de Bull. Sur le segment du milieu jusqu'au haut de gamme, nous avons une présence qui représente environ 15 à 25 % de ce marché. Sur la région Amériques, nous avons plus vocation à nous développer dans le sud - où nous avons déjà des filiales assez importantes - qu'à viser les Etats-Unis où le marché nous apparaît bouché par les compétiteurs locaux. Bien sûr, nous pouvons faire du business là-bas, mais ce n'est pas sur ce marché que porteront principalement nos efforts de développement.

Les calculateurs se dirigent vers toujours plus de parallélisme et Bull a inauguré au printemps dernier son Centre d'Excellence pour la Programmation Parallèle. Quels types d'applications visez-vous et peut-on déjà parler de premières retombées ?

Nous considérons l'utilisation des accélérateurs manycore Xeon Phi comme une tendance importante pour un certain nombre d'utilisateurs. De ce fait, nous souhaitons leur apporter une expertise susceptible de déboucher sur l'optimisation de leurs applications et sur ce plan, nous n'avons pas d'a priori. Il est cependant assez clair que certaines vont mieux se prêter au jeu que d'autres. Parfois, il vaut mieux repenser le problème plutôt que de partir d'un code dont on sait pertinemment qu'il ne présentera jamais un niveau de parallélisme suffisant pour une exécution optimale sur un cluster MIC.

Depuis une dizaine d'années, les systèmes et architectures HPC se sont diversifiés, et avec eux les modèles de programmation. Comment, dès lors, pérenniser les parcs applicatifs et assurer la portabilité des codes ?

C'est effectivement un enjeu que le SRA ETP4HPC a clairement identifié. Il nous semble aujourd'hui nécessaire de mettre en place un certain nombre d'API permettant d'optimiser les aspects applicatifs et les aspects exécution. Ces API, intervenant aux niveaux des modèles de programmation et des runtimes, vont vite devenir indispensables à une exploitation optimale des ressources. Il y a également un défi sur les I/O car les architectures I/O vont beaucoup évoluer. Le modèle mémoire / disques va sans doute être remis en cause à moyen terme, au moins sur les grands systèmes, en raison des différentiels de temps d'accès. Nous avons besoin de niveaux intermédiaires dans la hiérarchie du stockage, et de plus d'intelligence dans la façon dont les données sont traitées.

Concrètement, ces API devront par exemple permettre aux développeurs applicatifs de spécifier pourquoi et à quels moments ils accèdent aux données. Ajouter un peu de sémantique aide les fournisseurs de solutions middleware et matérielles à optimiser les

transports de données dans les architectures. Ces transports sont importants à la fois pour la performance et pour l'efficacité énergétique. Transporter les données dans un système HPC à très grande échelle consomme une partie non négligeable de l'enveloppe énergétique globale.

La gestion de la résilience est un autre exemple où l'émergence d'API est importante. Les développeurs applicatifs savent quelles données conserver et ce qui est nécessaire pour reprendre en cas de défaillance. Il faut avoir la capacité de le spécifier et d'implémenter des fonctions de résilience en rapport avec les différentes architectures qui vont se développer, en tenant compte notamment des nouveaux niveaux de stockage que nous évoquions.

Il y a clairement un gros effort à faire dans les années qui viennent pour définir ces API et pour que tous les intervenants concernés se mettent d'accord. Les développeurs applicatifs devront pouvoir utiliser ces API pour exprimer la sémantique de leurs codes et travailler sans avoir à tenir compte des modèles fins des machines. Si ces API tardent à apparaître, la communauté risque de perdre du temps, chacun allant dans sa propre direction.

Avec l'intégration massive des unités de calcul (CPU, GPU, APU) en perspective, considérez-vous la programmabilité comme un défi technique majeur ?

Il est certain que concevoir des applications fortement parallèles est un exercice intrinsèquement complexe. Le défi, de mon point de vue, est d'arriver à programmer en permettant l'expression de ce parallélisme massif. Pour faire le lien avec la question précédente, si on dispose d'API bien faites, on n'est plus obligé de se préoccuper de savoir si c'est un CPU ou un GPU qui calcule. Le runtime doit pouvoir aider et prendre les décisions en fonction des données et des ressources qui seront utilisées, adapter l'exécution de l'application à l'environnement matériel.

Pensez-vous, comme beaucoup d'observateurs, que l'efficacité énergétique soit l'obstacle majeur sur la route de l'Exascale ?

A l'évidence, l'efficacité énergétique est un enjeu majeur. Les roadmaps processeurs ne permettent pas aujourd'hui d'envisager l'exascale à 20 MW en 2020. Sur la simple puissance de calcul, on est encore très en deçà de ce qu'on cherche à obtenir. On doit également prendre en compte les mouvements de données, qui méritent eux aussi beaucoup d'optimisations pour que les solutions proposées aux utilisateurs fonctionnent dans des enveloppes énergétiques raisonnables.

Cela étant, il faut comprendre qu'avec 10^{18} opérations par seconde sur des processeurs cadencés à quelques GHz, on atteint quasiment le milliard d'opérations en parallèle. Il va nous falloir des méthodes de résolution qui facilitent les calculs en simultané avec le moins de synchronisations possible. Cette maîtrise du parallélisme est sans doute un défi aussi important que celui de l'efficacité énergétique.

Pour conclure, quelles grandes ruptures technologiques voyez-vous se profiler d'ici 2020 ?

Je pencherais plutôt pour des évolutions fortes que pour de véritables ruptures. J'ai quelques doutes, personnellement, sur l'émergence prochaine de superordinateurs quantiques, de machines au graphène, d'architectures neuro-mimétiques pour le calcul, etc. Nous sommes encore aujourd'hui sur des technologies logicielles et matérielles - notamment le CMOS - qui permettent de continuer à avancer. Or, les alternatives que nous venons de citer sont loin d'avoir le même degré de maturité. Sans doute arriveront-elles, mais pas en 2020... ■

CARRI Systems

Architecte technologique depuis 1992

Intel Technology Provider Platinum 2013

NVIDIA PARTNERFORCE TESLA PREFERRED PROVIDER

Concepteur d'Architectures Rationnelles pour la Recherche et l'Industrie

«Créée en 1992 par des passionnés, CARRI Systems est le dernier fabricant français informatique haut de gamme sur-mesure. Basée à Noisy-le-Sec, près de PARIS, la société est, depuis quelques années, devenu un acteur majeur sur le marché du GPU Computing et du calcul intensif. CARRI Systems est membre du Pôle européen de compétence en simulation numérique haute performance (Ter@tec) depuis 2011.»

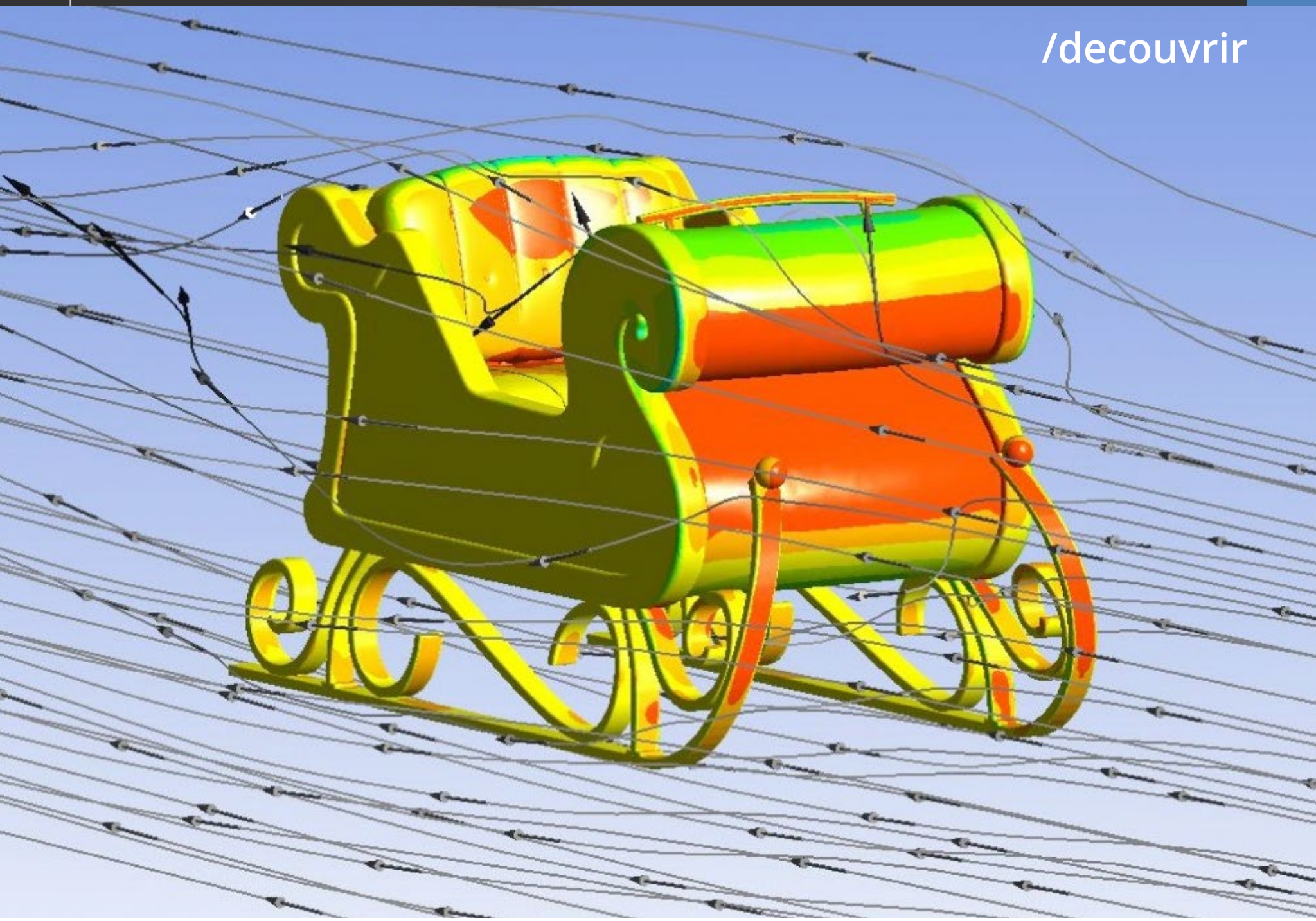
Franck Darmon
CEO CARRI Systems

intel NVIDIA AMD

FABRIQUÉ EN FRANCE

Découvrez toutes nos solutions XLR sur

www.carri.com



SIMULER LA TURBULENCE : DES SOLUTIONS POUR L'INDUSTRIE

Longtemps synonymes d'imprécision dans la simulation des écoulements fluidiques, les phénomènes de turbulence sont aujourd'hui à peu près maîtrisés par les grandes solutions de CFD. Reste à choisir, parmi les modèles qu'elles proposent, le mieux adapté aux besoins applicatifs et aux ressources de calcul à disposition…

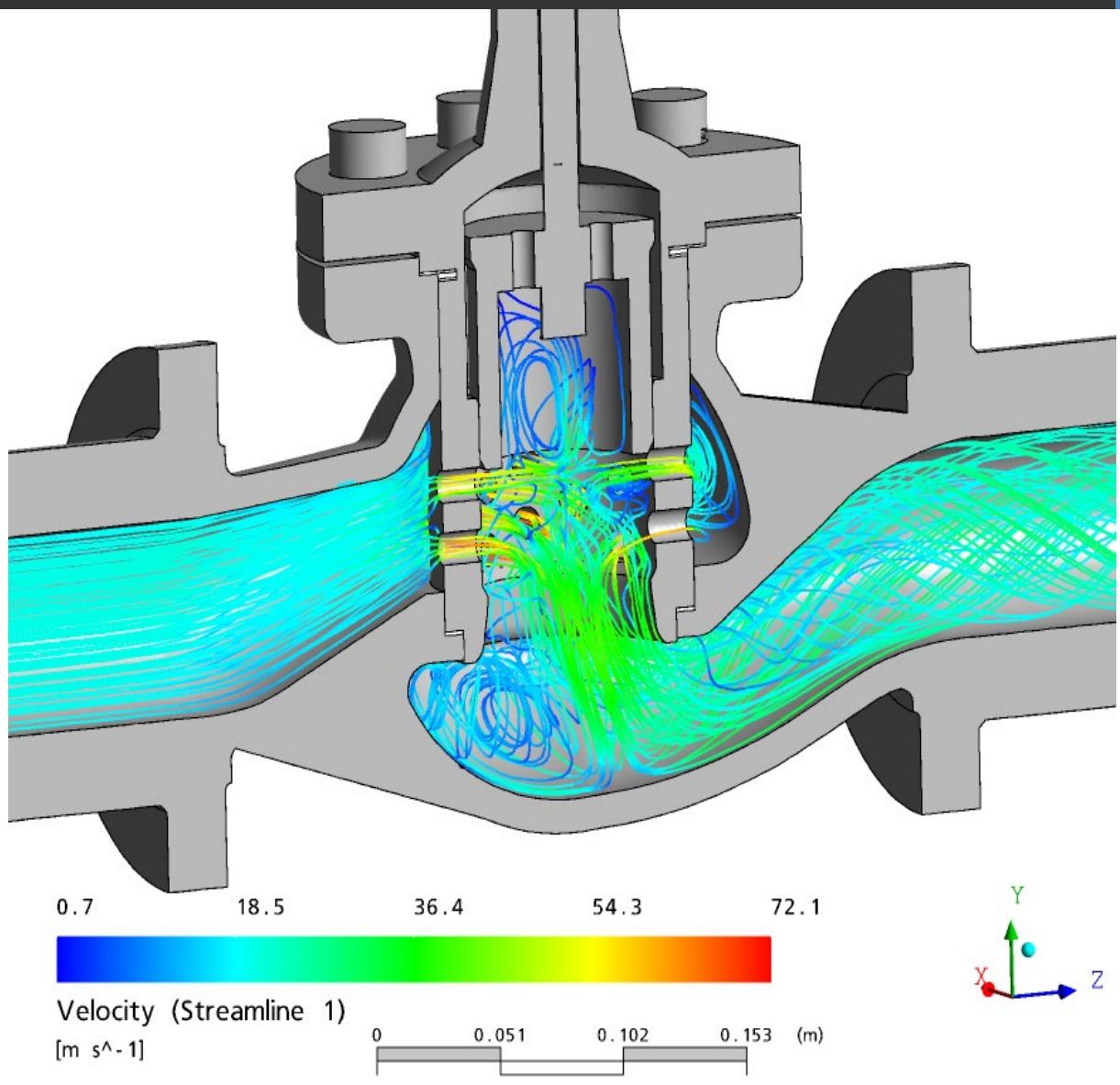
GILLES EGGENSPIELER, PhD

La turbulence joue un rôle important dans la grande majorité des applications industrielles d'écoulement des fluides. C'est une problématique multiéchelle classique en mécanique des fluides dans la mesure où les échelles présentes dans un écoulement turbulent interagissent les unes avec les autres. Pour connaître la performance d'un système avec la précision requise, il faut pouvoir prévoir au plus juste une multiplicité de facteurs : propriétés aérodynamiques, caractéristiques de transfert thermique, performances des mélanges, etc. Modéliser la turbulence avec suffisamment de fidélité est tout aussi important dans les contextes de la mécanique des fluides numérique (CFD / Computational Fluid Dynamics). Les ressources de calcul dont on

* Sr. Product Manager, ANSYS.

dispose aujourd'hui sont généralement insuffisantes pour permettre la résolution numérique (simulation numérique directe ou DNS) de toutes les échelles d'écoulements turbulents présentes dans les applications industrielles. Il existe néanmoins des équations capables de simplifier le problème et de donner des informations justes sur la turbulence d'un écoulement et ses effets sur la performance d'un produit.

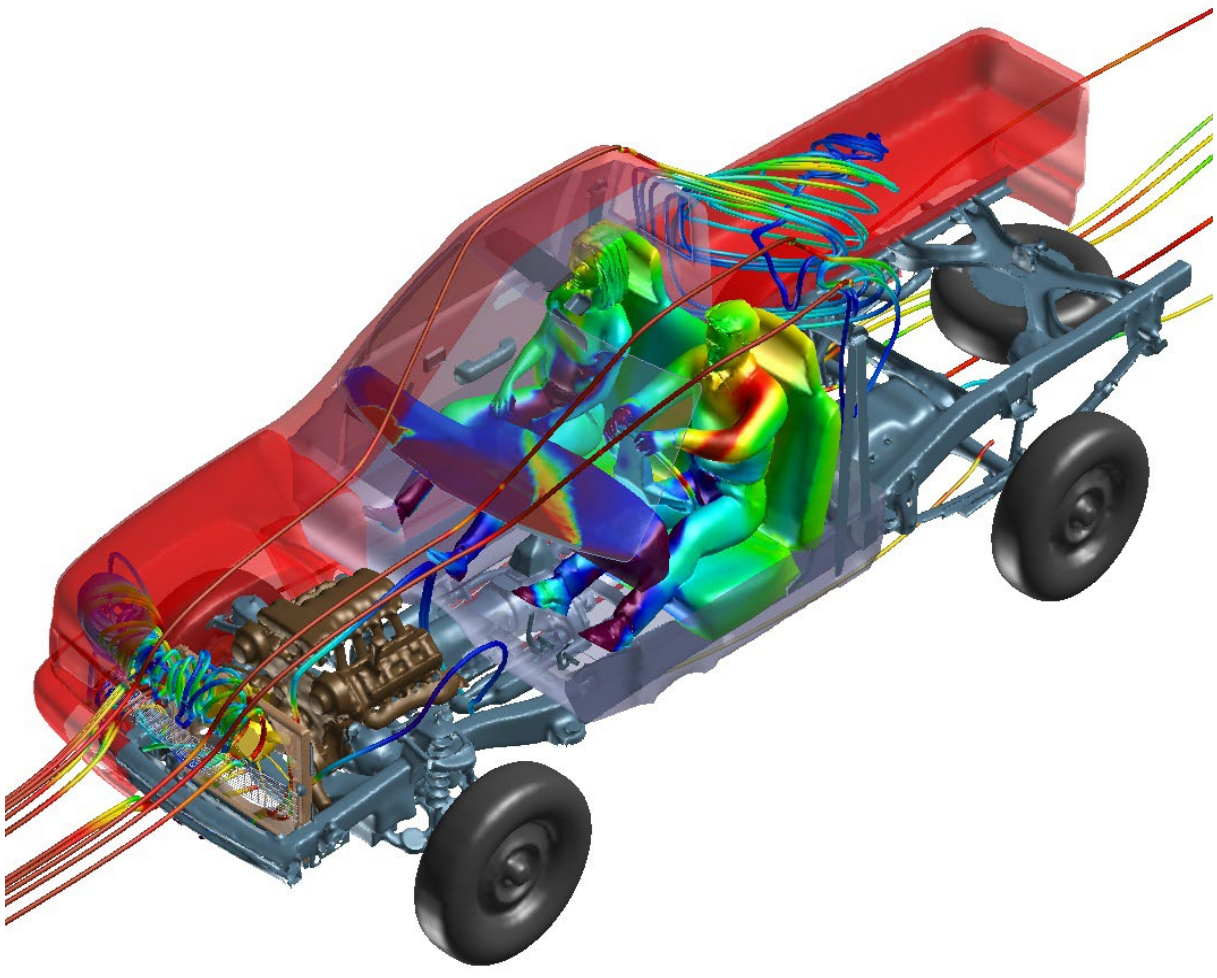
La turbulence est un phénomène si complexe qu'il n'existe pas (encore) de formulation universelle pour le résoudre. Face à la multitude des modèles disponibles, les fournisseurs d'applications de CFD doivent intégrer ceux qui leur paraissent les plus pertinents, de façon à répondre à la majorité des besoins des utilisateurs.



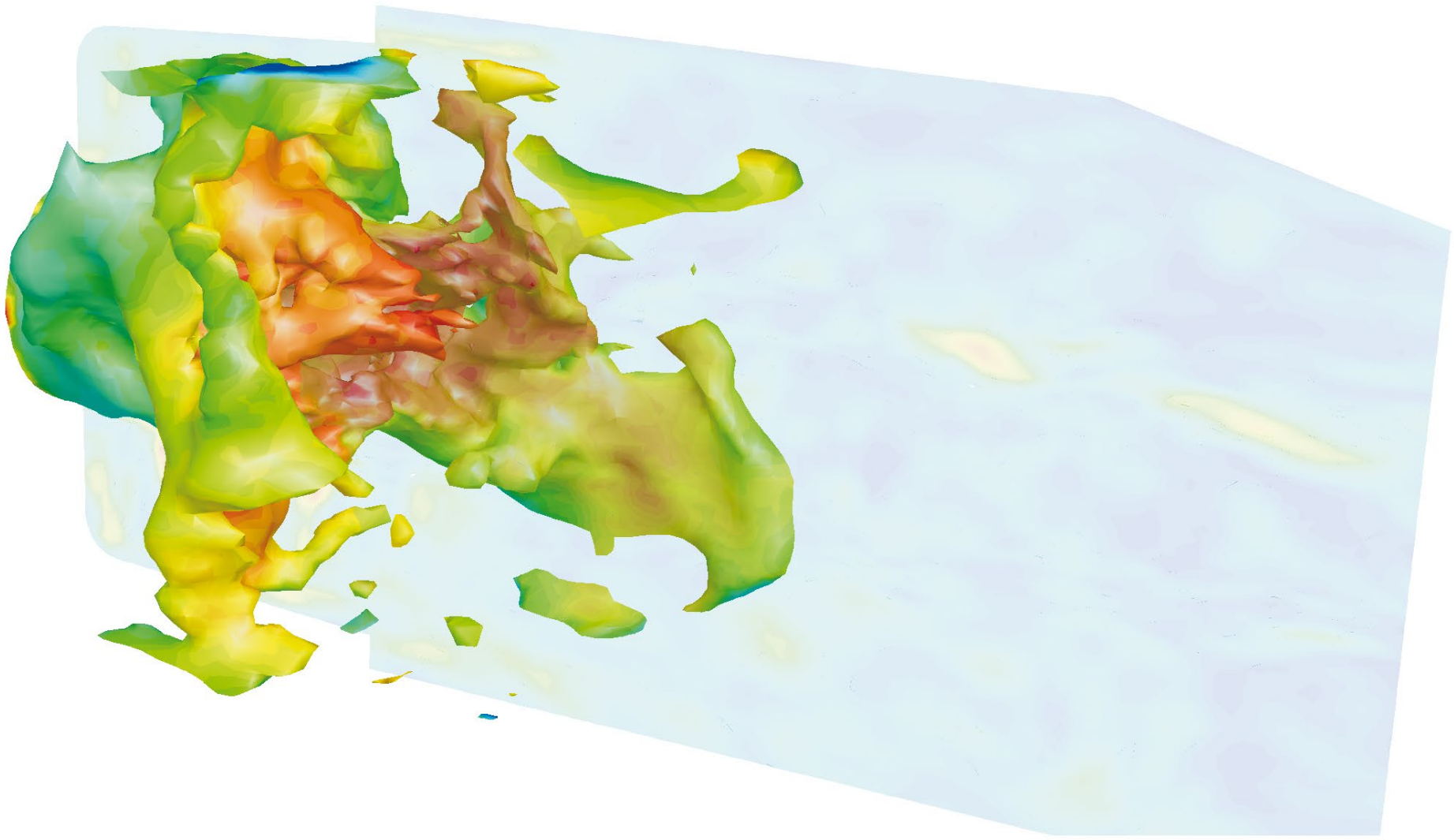
Prédire le fonctionnement exact d'un système de valve complexe nécessite de simuler les caractéristiques d'écoulement dans la conduite principale et dans les petites ouvertures. Les effets à proximité des parois doivent également être pris en compte, là où les turbulences les plus complexes sont initiées.

Ecoulements multiéchelles, tridimmensionnels, instables

Certains spécialistes du secteur vont parfois plus loin en proposant des recommandations quant au modèle à utiliser face à tel problème de turbulence spécifique. Ils contribuent aussi aux progrès de la physique par leurs efforts de recherche, de test et de validation de nouveaux modèles mixant plusieurs approches de représentation de la turbulence (modèles hybrides RANS/LES par exemple) ou permettant la prise en compte de phénomènes de transition laminaire-turbulent - modèles dont la vocation est d'offrir le meilleur compromis entre exactitude des résultats numériques et ressources nécessaires pour le calcul.



Les constructeurs automobiles ont recours à la CFD pour calculer précisément la turbulence d'écoulement du chauffage et du refroidissement des véhicules particuliers.

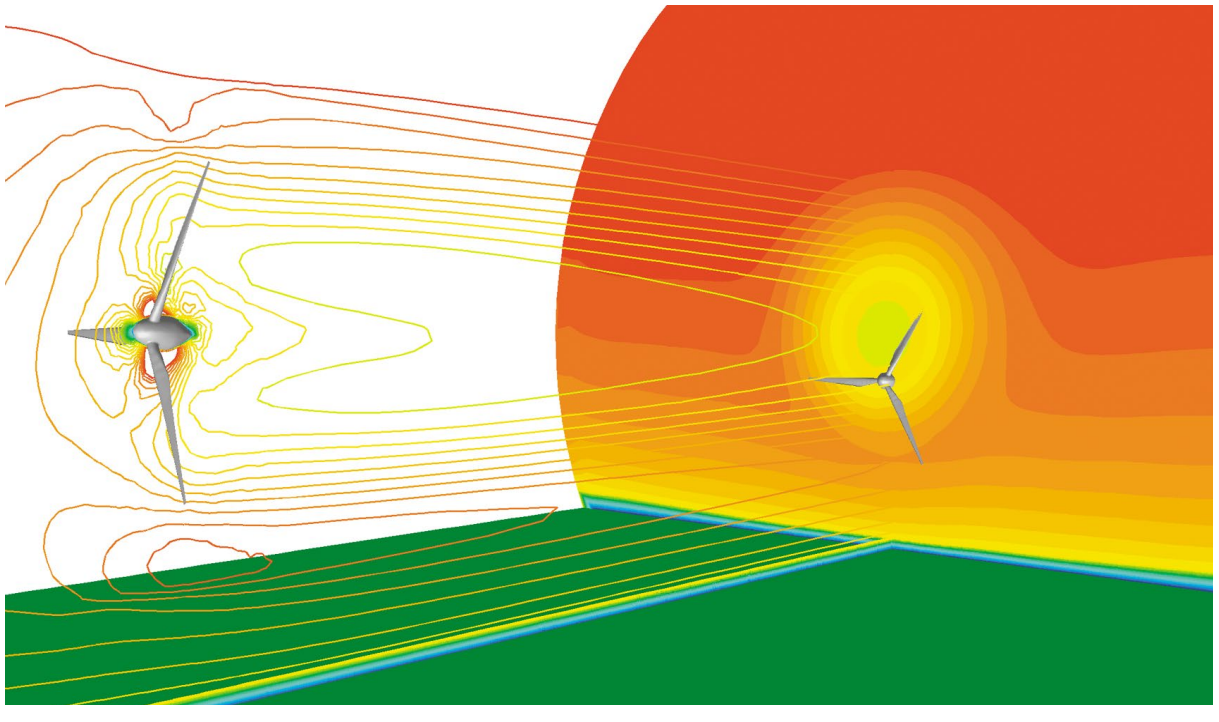


Les modèles LES permettent d'obtenir des résultats plus fins pour certains types d'applications. Dans le cas de cette chambre de combustion, il faut simuler avec précision les grandes échelles pour pouvoir prédire de façon opératoire la position de la flamme, le dégagement de chaleur et les émissions de polluants.

En ingénierie, les écoulements sont le plus souvent turbulents et, par nature, multiéchelles, tridimensionnels et instables. Les écoulements de fluides les plus difficiles à modéliser se rencontrent notamment dans les secteurs industriels suivants :

Aéronautique - La modélisation précise de la turbulence intervient dans les problématiques aéronautiques les plus critiques, notamment lorsqu'il s'agit d'optimiser le rapport portance-traînée d'une aile afin de garantir la sécurité de vol de l'avion pour une consommation de carburant la plus faible possible. La modélisation de la turbulence intervient également dans la conception des moteurs, en particulier pour limiter le bruit à l'extérieur et dans la cabine.

Oil & gas - Sachant que toute turbulence entraîne une perte d'énergie dans l'écoulement, la modélisation devient critique lorsque l'on doit faire circuler des fluides sur de longues distances. Il faut par exemple estimer précisément la déperdition d'énergie pour déterminer la meilleure distance entre deux stations de pompage. Si la distance est trop grande, l'écoulement perd de la vitesse voire s'arrête ; si la distance est trop

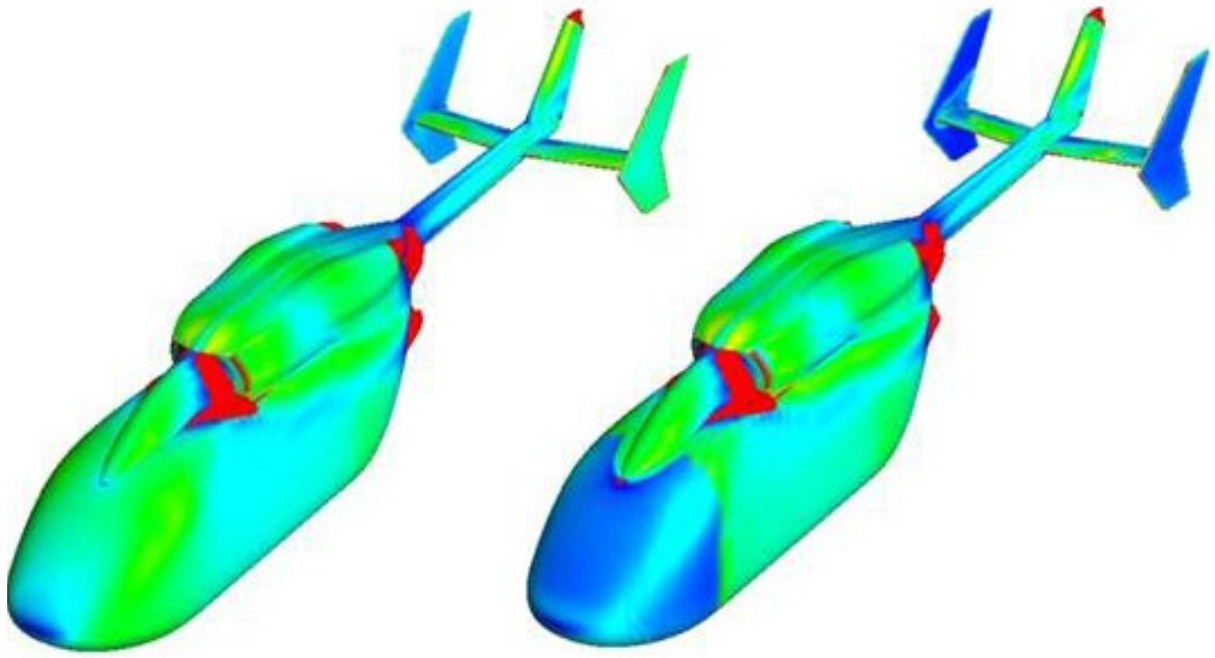


Seul un modèle de turbulence précis peut permettre de prédire les contours de vélocité dans le sillage d'une éolienne. Cette dimension est critique lorsqu'on étudie l'effet de la traînée d'une rangée d'éolienne sur celle située en aval.

courte, le système de pompage n'est pas optimal, avec ce que cela suppose de gaspillages (coûts, maintenance...). Pouvoir prédire précisément les comportements d'écoulement et la chute de pression induite par la turbulence permet de concevoir des systèmes de transport de pétrole et de gaz plus rentables et plus efficaces.

Automobile - L'aérodynamique automobile est faite de compromis, surtout quand il s'agit de concilier style de la carrosserie et contraintes de Cx. Comparer l'effet de la traînée de différentes propositions de conception permet de sélectionner le design le plus convaincant mais aussi celui qui consommera le moins de carburant. La turbulence dictant la performance aérodynamique d'un design, faire le bon choix suppose de pouvoir compter sur des modèles de turbulence précis, testés et validés.

Chimie et *drug design* - Les opérations de mixage sont cruciales dans la plupart des procédés chimiques et pharmaceutiques. Pour obtenir des résultats optimaux, il faut parvenir à des mélanges homogènes, et cela dans le temps le plus court possible. Les fabricants d'équipements de mixage utilisent la CFD pour comparer différents modèles de malaxeurs et de réservoirs, et ainsi déterminer la configuration optimale au regard des conditions opératoires. La turbulence jouant un rôle clé dans l'opération de mélange, le choix du modèle turbulent conditionne le nombre d'itérations nécessaires pour obtenir le bon design.



Prédiction du coefficient de résistance d'un hélicoptère. A gauche, les résultats obtenus au moyen d'un modèle de turbulence RANS surévaluent la traînée effective du fuselage. A droite, un modèle de transition donne des résultats nettement plus précis.

Energies vertes - Pour les compagnies d'électricité qui exploitent des éoliennes, l'objectif est que chaque ferme produise le maximum d'énergie au moindre coût. Le sillage d'une éolienne a de fortes incidences sur les turbines situées sous le vent - pertes de rendement et raccourcissement de la longévité, notamment. Les ingénieurs qui conçoivent l'implantation des fermes et le positionnement des éoliennes ont donc besoin de méthodes avancées pour le calcul de ces effets de sillage et de terrain. Prévoir précisément ces phénomènes au moyen de simulations numériques permet d'obtenir des rendements idéaux.

Modéliser la turbulence pour les écoulements complexes

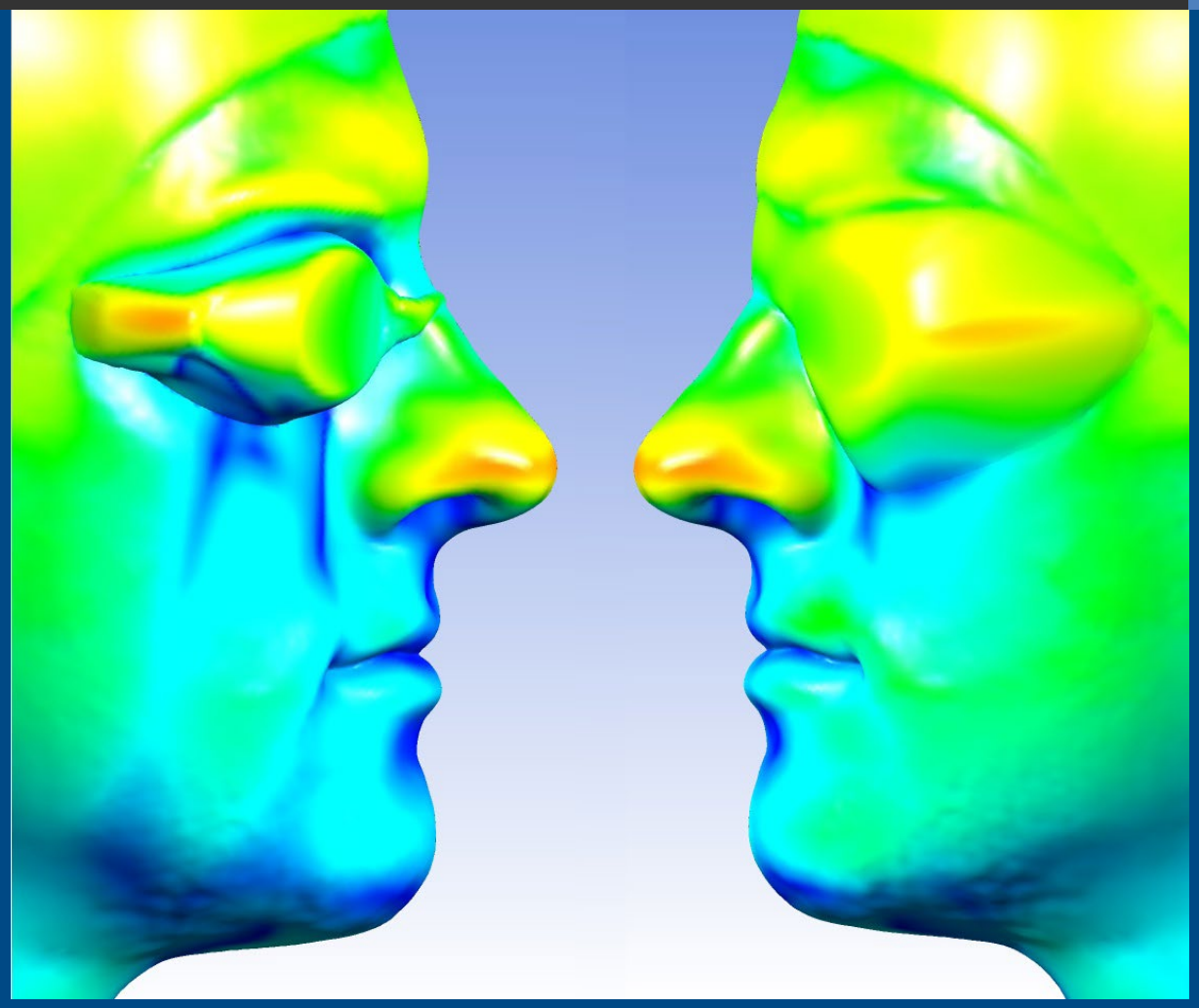
S'il n'existe pas de modèle ou d'approche unique de la modélisation pour tous les écoulements turbulents, différents types de modèles ont été développés au cours des dernières décennies qui permettent d'obtenir des résultats précis dans les applications acoustiques, de combustion et d'autres applications similaires, où la résolution d'une partie des échelles turbulentes s'avère nécessaire pour obtenir des résultats fins. Voici une brève présentation des principaux :

Modèles *steady-state* - Les modèles stationnaires ou RANS (Reynolds-Averaged Navier–Stokes) simplifient le problème de la modélisation des écoulements turbulents et résolvent des grandeurs moyennées de l'écoulement (vitesse, pression, densité et température). L'approche stationnaire calcule des quantités d'écoulement moyennes, sans aucun effort de résolution des phénomène de tourbillons dans le temps et dans l'espace.

Ces méthodes RANS ne consomment donc que peu de ressources machine. Pour les ingénieurs qui sélectionnent les bons modèles et les appliquent correctement, cette solution permet de prévoir l'effet de turbulence à moindre coût, c'est-à-dire sans devoir explicitement capturer l'ensemble des

Quand Speedo modélise les turbulences complexes

Lors des compétitions olympiques de natation, l'écart entre l'or et l'argent se mesure en millisecondes, de sorte que d'infimes améliorations peuvent faire la différence. C'est pourquoi Speedo, équipementier bien connu dans cette discipline, utilise les logiciels de CFD pour éliminer toute imperfection susceptible d'augmenter la résistance à l'avancement et de ralentir le nageur.

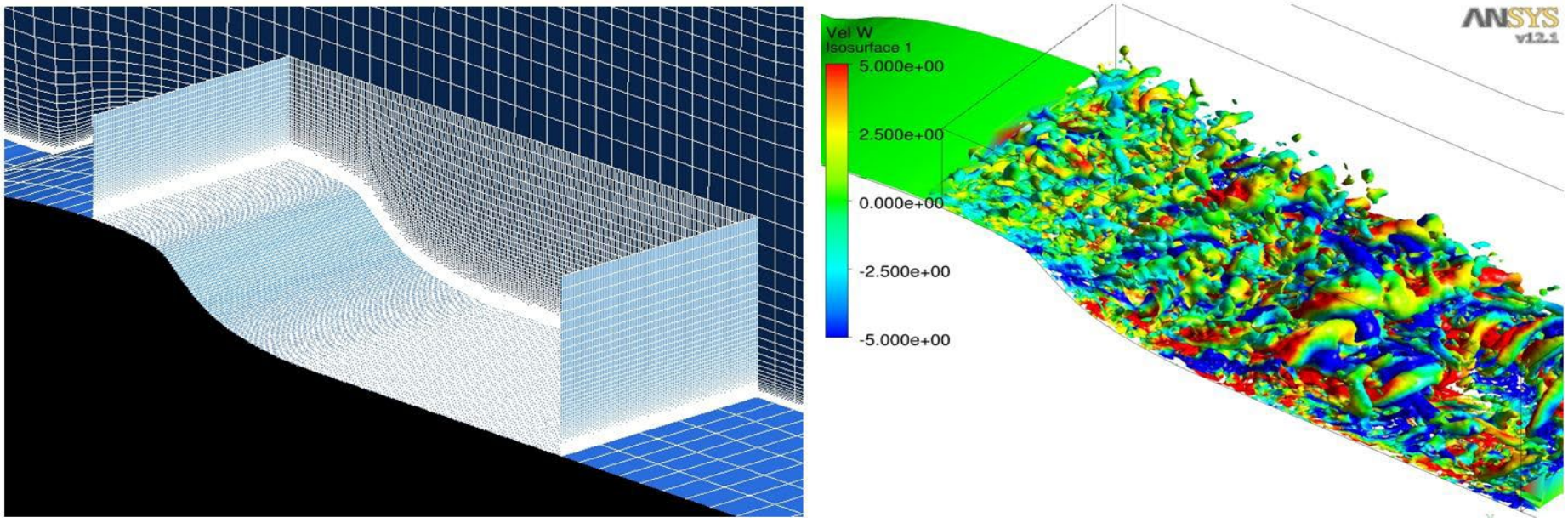


Les ingénieurs de Speedo Aqualab travaillent évidemment sur les trois composantes de l'équipement du nageur - maillot, lunettes et bonnet. La difficulté, c'est que la combinaison des conditions d'écoulement autour de ces éléments et du comportement physique du nageur crée un environnement de turbulence extrêmement complexe.

Il leur faut donc des modèles de turbulence très précis pour aboutir aux designs qui donneront un avantage compétitif au nageur, avant même que la course à proprement parler ne démarre. Les spécialistes de la natation ont ainsi pu observer une amélioration continue des équipements des nageurs, notamment ceux que Speedo sponsorise. C'est le cas des maillots et des bonnets qui font littéralement corps avec l'athlète et diminuent assez sensiblement les perturbations du champ d'écoulement. De même, les lunettes sont désormais conçues pour épouser les caractéristiques physiques du visage du sportif. Chronomètre en main, les résultats sont bien au rendez-vous : les lunettes conçues en modélisation CFD permettent par exemple d'améliorer les performances de l'ordre de 2 %.

échelles des écoulements. Les techniques RANS sont très précises dans la grande majorité des applications industrielles. Il faut toutefois recourir à des modèles plus sophistiqués, donc plus gourmands en ressources, pour celles qui nécessitent la résolution d'une plage d'échelles turbulentes.

Modèles LES (Large-Eddy Simulation) et hybrides - Les modèles LES de simulation aux grandes échelles reposent sur la résolution des grandes structures de turbulence dans le temps et dans l'espace, et simulent l'influence des petites structures de turbulence. Plus sophistiqués, ces modèles s'avèrent d'un coût machine supérieur, d'environ un à deux ordres de grandeur, par rapport aux modèles RANS. A mesure que les tourbillons diminuent de taille à proximité des parois, utiliser ces modèles dans le processus de développement revient donc assez cher, ce qui pose des problèmes dans les

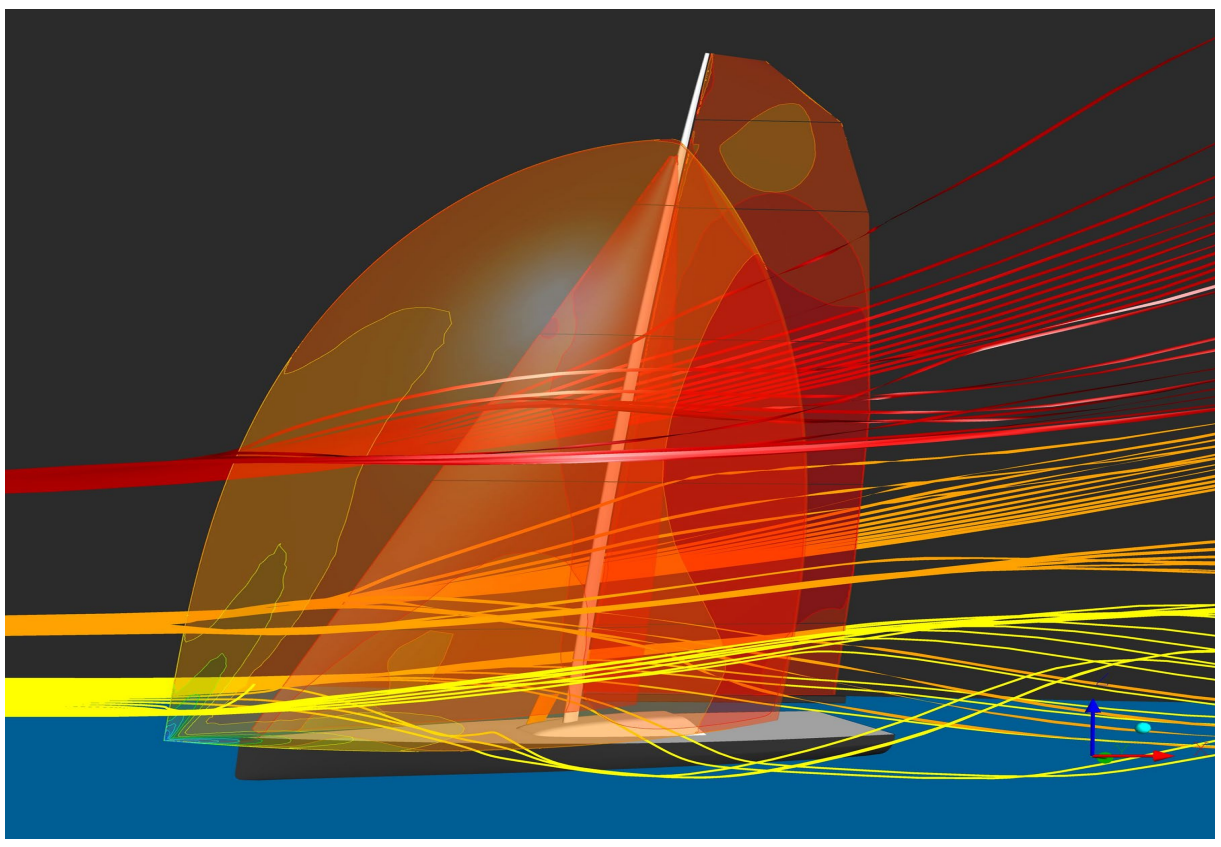


En pointe dans le développement de modèles hybrides, ANSYS a récemment mis au point WMLES (Wall-Modeled LES) et ELES (Embedded LES). WMLES permet la simulation des écoulements à proximité des parois sur des infrastructures de calcul plus légères qu'avec la modélisation LES classique. ELES offre quant à lui plus de flexibilité encore, avec une zone LES embarquée et un plus vaste domaine steady-state avec simulation RANS. Cette double image montre les structures tourbillonnaires dans le sillage du stabilisateur de flamme d'un corps non profilé Volvo, obtenues par simulation DDES (Delayed Detached Eddy Simulation).

simulations où, précisément, les effets de parois impactent le design et la performance du produit final. On peut alors envisager des modèles hybrides, DES (Detached-Eddy Simulation) et SAS (Scale-Adaptive Simulation) qui appliquent une combinaison de traitements *steady-state* et LES aux couches limites.

Modèles de transition - Concernant les écoulements autour des pales d'éoliennes, des ailes d'avion, etc., on observe souvent que des couches limites laminaires en amont se transforment progressivement, en aval, en écoulement turbulent. Il faut comprendre ce phénomène pour prévoir la portance des ailes ou la performance d'un compresseur par exemple. Il existe désormais une nouvelle catégorie de modèles pour analyser ces phénomènes physiques et prévoir le comportement d'un écoulement sur une surface qui passe du régime laminaire au régime de turbulence. Ces modèles, adaptés aux écoulements de transition, permettent d'envisager nombre de nouvelles applications pour les logiciels CFD.

De ce qui précède, il faut conclure que la turbulence intervient dans la plupart des problèmes d'écoulement et qu'elle est le facteur qui limite le plus la précision de simulation. S'il n'existe pas d'approche universelle pour la modélisation de tous les écoulements turbulents à toutes les échelles, différents types de modèles sont aujourd'hui disponibles pour traiter l'immense majorité des besoins industriels. La principale difficulté consiste donc à choisir dès le départ le modèle le mieux adapté à l'application étudiée et à l'infrastructure de calcul dont on dispose... ■



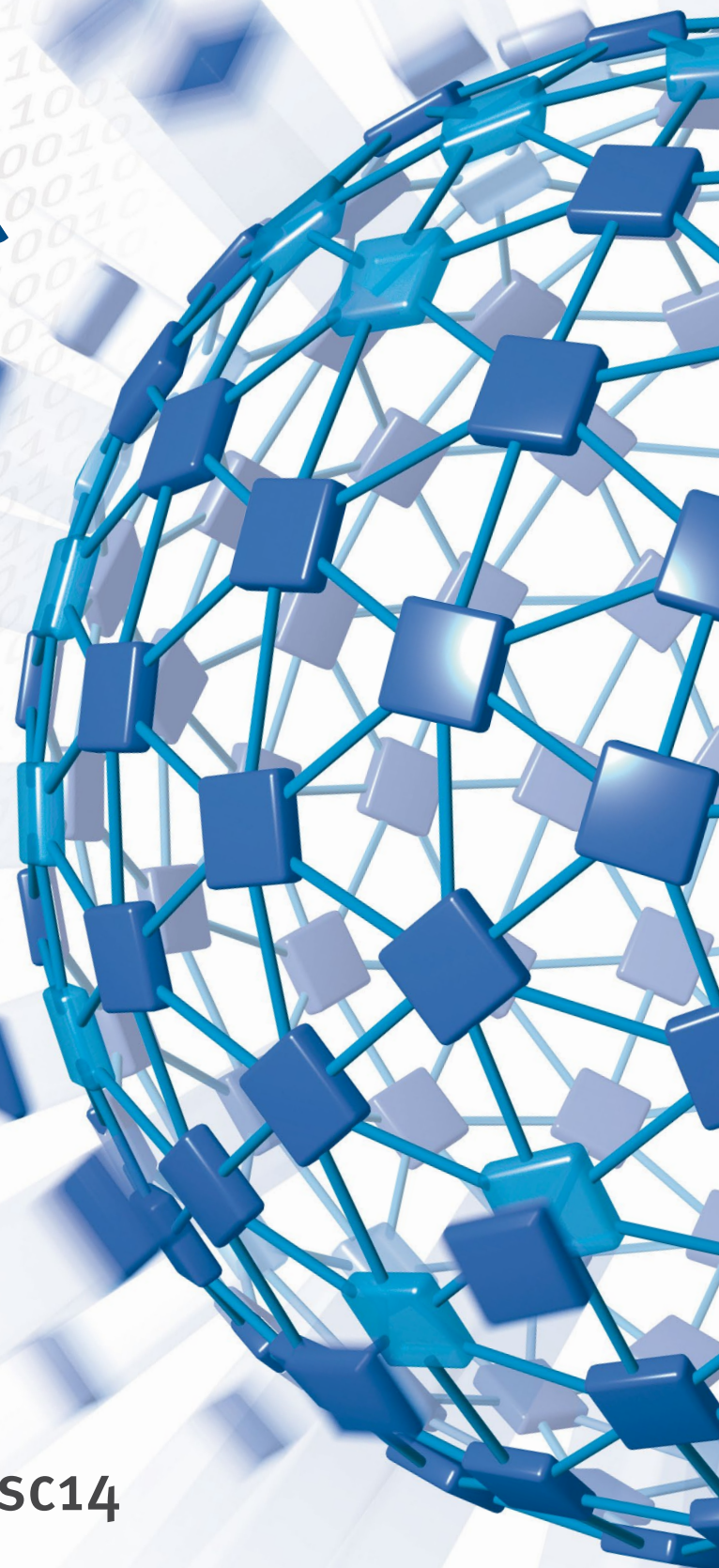
Modéliser précisément la turbulence est aujourd'hui indispensable pour concevoir des voiliers de course à très très hautes performances.

INTERNATIONAL
SUPERCOMPUTING CONFERENCE

ISC'14 THE HPC EVENT

June 22 – 26, 2014, Leipzig, Germany

Join the Global Supercomputing Community



www.isc-events.com/isc14

/l'état_de_l'art

L'ECOSYSTEME DES BIBLIOTHEQUES POUR LE CALCUL SCIENTIFIQUE

Utiliser ou ne pas utiliser les librairies scientifiques? Cette question, cruciale pour le chercheur ou l'ingénieur en charge d'une application ciblée, semble avoir perdu de sa pertinence. Les enjeux économiques et sociétaux, ajoutés à la qualité globale des outils disponibles en Open source, tendent aujourd'hui à favoriser l'innovation collective. Voici pourquoi...

PHILIPPE D'ANFRAY[1], MICHEL KERN[2], THIÊN-HIÊP LÊ[3], FABIEN NICOLAS[4]

Les bibliothèques logicielles scientifiques sont à la fois des composants majeurs des grands logiciels de simulation et des portes d'accès au calcul haute performance (HPC). On trouve en Open source des packagings performants, validés et maintenus, qui permettent de "cristalliser" nombre de projets de recherche dans divers domaines (aérodynamique, climat, matériaux, chimie...) et garantissent le partage de savoir-faire par des acteurs d'horizons très variés. Le gros avantage des bibliothèques, notamment dans le calcul, c'est qu'elles déchargent du codage de fonctionnalités ancillaires (produit scalaire, matriciel, racine carrée d'un vecteur...) et, dès lors qu'elles sont un tant soit peu optimisées, améliorent considérablement les performances des codes. Ne pas y recourir permet certes de conserver la maîtrise d'œuvre du développement applicatif, mais rend la tâche significativement plus

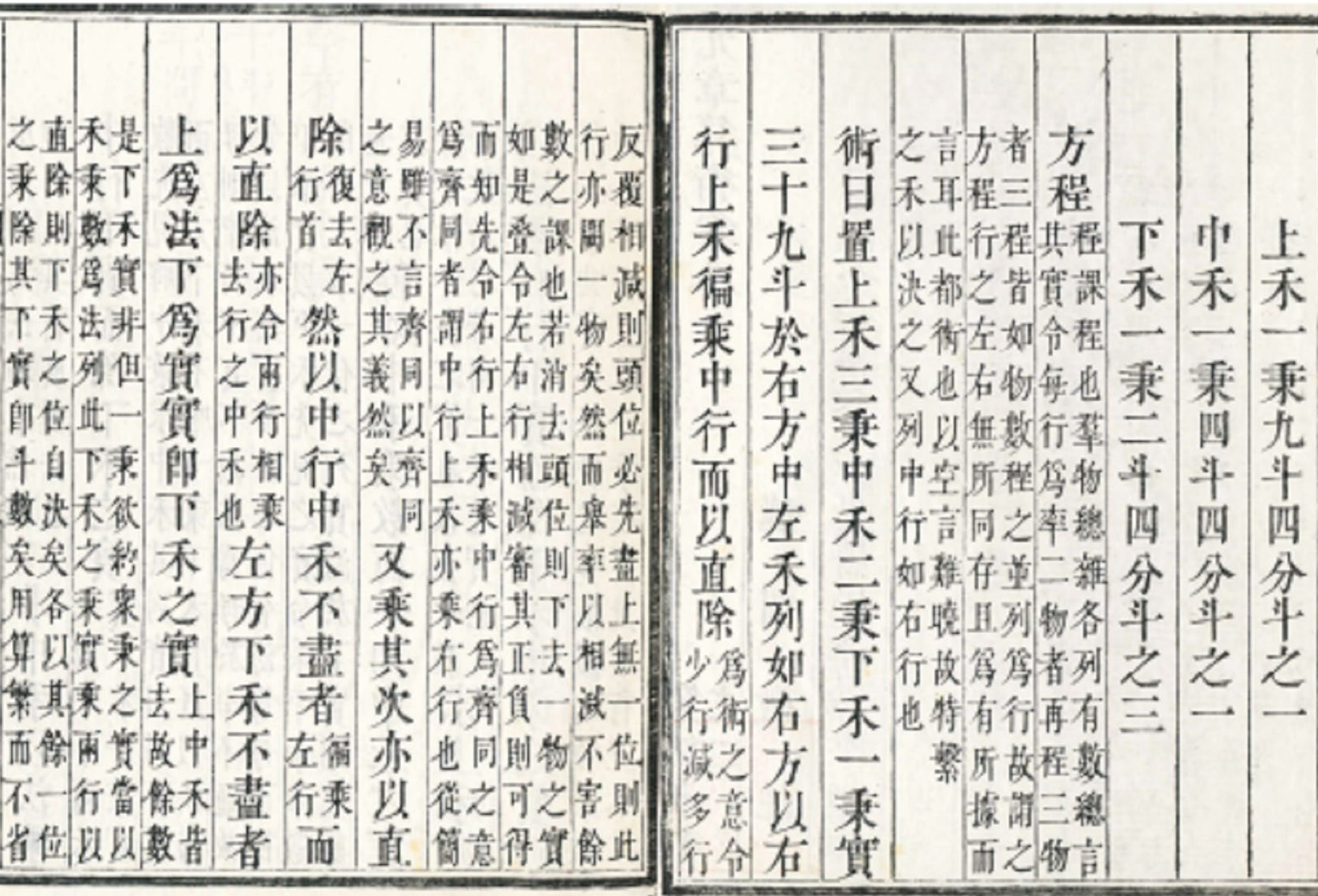
[1] CEA ; [2] INRIA ; [3] ONERA ; [4] UMAPS.

lourde. Le **séminaire** Aristote qui s'est tenu en mai 2013 à l'Ecole polytechnique a permis de faire la lumière sur de nombreuses réalisations et de mieux connaître l'écosystème y afférent. Cet article a vocation à en résumer les principaux échanges.

Travailler confortablement avec l'algèbre linéaire

L'algèbre linéaire est omniprésente dans les simulations, la visualisation, les jeux vidéo... Dans les mathématiques, c'est la branche qui formalise la théorie des systèmes d'équations linéaires ou des matrices issus de la discrétisation spatiale et de l'intégration temporelle des équations aux dérivées partielles modélisant l'application considérée (prévision du climat, modélisation de l'environnement, construction aéronautique etc.). Il existe plusieurs types d'outils pour travailler avec l'algèbre linéaire, mais aucun n'est pleinement satisfaisant. Des solutions comme MATLAB, par exemple, sont certes très *user-friendly* mais elles se cantonnent aux mathématiques pures. Quant aux bibliothèques de calcul intensif, elles sont assurément plus optimisées, mais demeurent très spécialisées. C'est précisément la raison pour laquelle Gaël Guennebaud (INRIA) a développé **Eigen**, une solution intermédiaire qui exploite les avantages des deux approches et peut s'utiliser à la fois sous MATLAB et en contexte HPC. Codé en pur C++ "moderne" (c'est-à-dire en *templates*), Eigen n'a pas besoin d'être configuré ni compilé. Il comprend également un certain nombre de fonctions et de composants - décomposition matricielle, géométrie, matrices creuses et solveurs - le tout unifié au sein d'une seule API.

Algorithme de Gauß : Extrait du chapitre 8 de Jiu Zhang Suan Shu (ca. 100 BC).



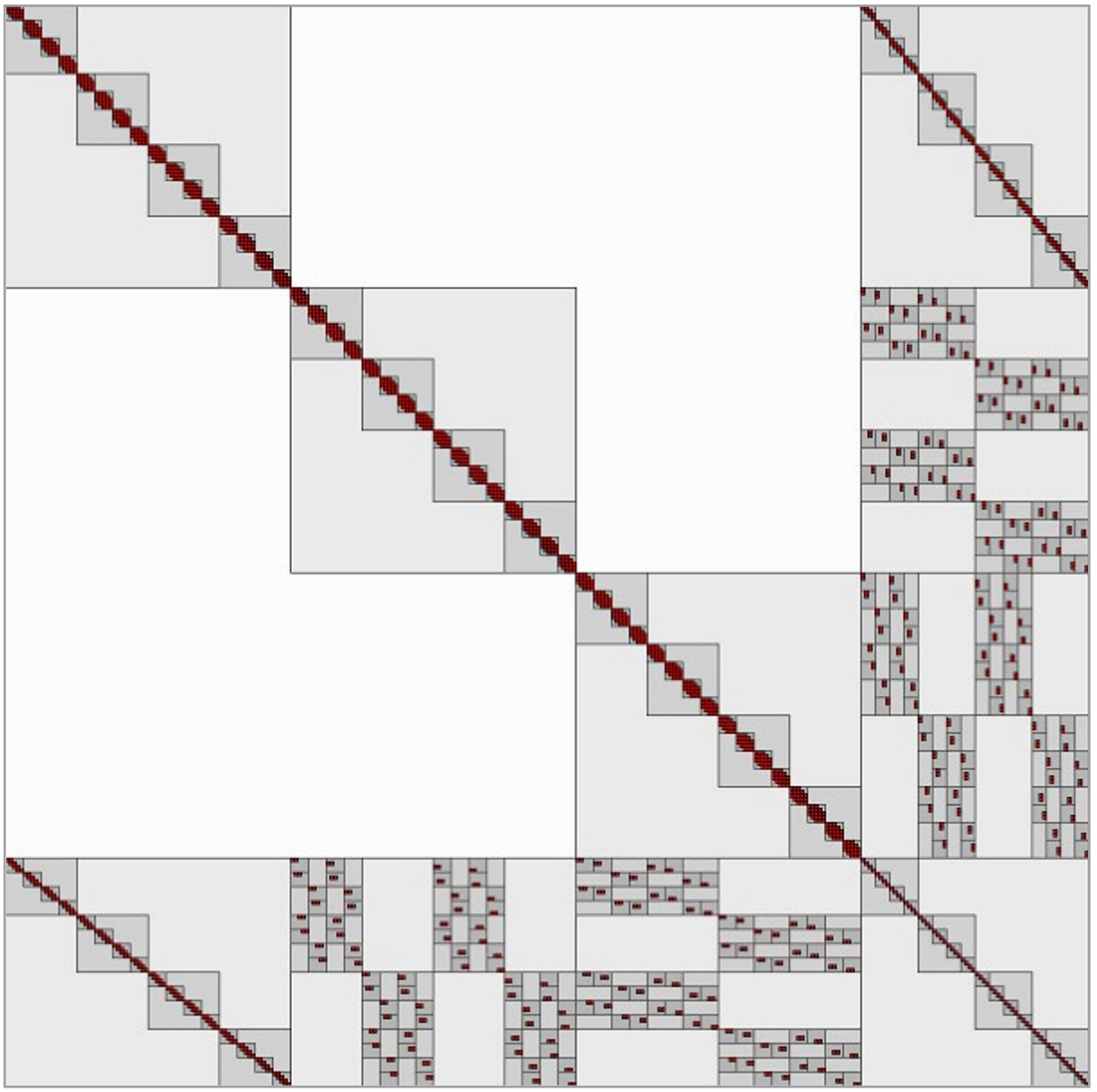
上禾一秉九斗四分斗之一
中禾一秉四斗四分斗之一
下禾一秉二斗四分斗之三

方程 程課程也羣物總雜各列有數總言其實令每行爲率二物者再程三物者三程皆如物數程之並列爲行故謂之方程行之左右無所同存且爲有所據而言耳此都術也以空言難曉故特繫之禾以決之又列中行如右行也

術曰置上禾三秉中禾二秉下禾一秉實三十九斗於右方中左禾列如右方以右行上禾徧乘中行而以直除爲術之少行減多行反覆相減則頭位必先盡上無一位則此行亦闕一物矣然而皋率以相減不害餘數之課也若消去頭位則下去一物之實如是叠令左右行相減審其正負則可得而知先令右行上禾秉中行爲齊同之意爲齊同者謂中行上禾亦乘右行也從簡易雖不言齊同以齊同之意觀之其義然矣 又乘其次亦以直除復去左行首 然以中行中禾不盡者徧乘左行而以直除亦令兩行相乘去行之中禾也左方下禾不盡者上爲法下爲實實即下禾之實上中禾皆去故餘數是下禾實非但一秉欲約衆秉之實當以禾秉數爲法列此下禾之秉實秉兩行以直除則下禾之位自決矣各以其餘一位之秉除其下實即斗數矣用算繁而不省

Reste que, comme le soulève François-Xavier Roux (ONERA), dans les systèmes implicites de très grande taille et mal conditionnés - comme ceux rencontrés en électromagnétisme - la complexité reste importante. Pour y faire face, il propose la méthode **FETI** (Finite Element Tearing and Interconnecting) qui consiste à décomposer l'ensemble en sous-domaines grâce aux éléments finis. Dans chaque domaine, les seuls prérequis sont la matrice de rigidité, les conditions aux limites locales issues des conditions aux limites globales, les contraintes et les interfaces. Pour appliquer cette méthode sur un objet existant, il faut mailler, autrement dit découper l'ensemble, soit "à la main", soit par un partitionneur de maillage tel **Scotch**. Développé par François Pellegrini (INRIA), Scotch est basé sur les graphes, outils primordiaux pour l'analyse. Constitués simplement de points reliés entre eux par des arrêtes, les graphes servent dans de nombreux applications - de l'analyse de réseaux routiers à la répartition de communications entre serveurs.

Les deux principaux problèmes étant la renumérotation de matrices creuses et la décomposition de domaines pour les méthodes itératives, Scotch a deux objectifs : *primo*, permettre le partitionnement-arête et le partitionnement-sommet ; *secundo* gérer des graphes de plus d'un milliard de sommets distribués sur des milliers de processeurs. Ces problèmes ont été globalement résolus à Livermore, même si le modèle "craque un peu aux membrures". Il s'agit donc à présent de pousser le système plus loin. L'horizon actuel est de 10^{15} sommets distribués sur un million de processeurs, un volume qui pose la question de la gestion des architectures massivement non uniformes (à la fois comme cibles mais aussi en redistribution à la volée) et de la hiérarchie sur la machine même où s'exécute Scotch. On s'oriente vers des algorithmes asynchrones pour relâcher les contraintes tandis qu'un certain nombre d'entre eux ont été reformulés pour le contexte multithreading.



Matrices creuses très structurées (blocs multiniveaux), typiques des problèmes de neutroniques et prises en charge par la bibliothèque Legolas++.

Des bibliothèques plus spécifiques

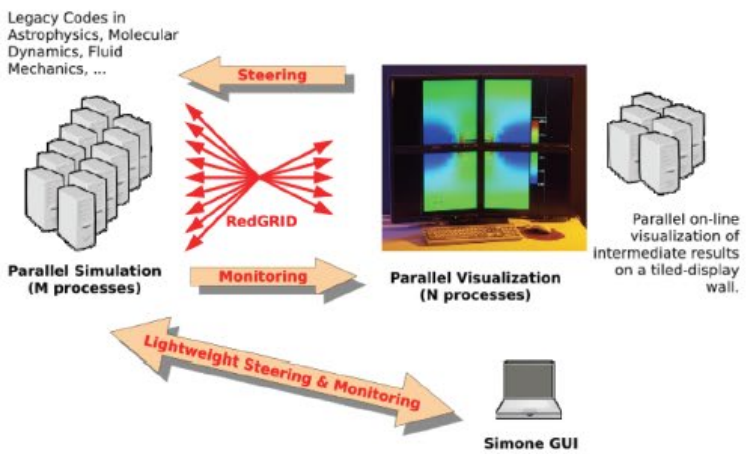
À EDF R&D, Laurent Plagne travaille sur un type de problème très particulier : les matrices creuses structurées par blocs multiniveaux. Ces objets exotiques sont engendrés par des problèmes de neutroniques et consistent en des matrices très creuses, dont on peut exprimer le *pattern* en peu d'éléments. Pour les décrire, il faut recourir à **Legolas++**, un outil qui construit des *patterns* complexes à partir de *patterns* de base, allant même jusqu'à reconstruire des matrices non stockées.

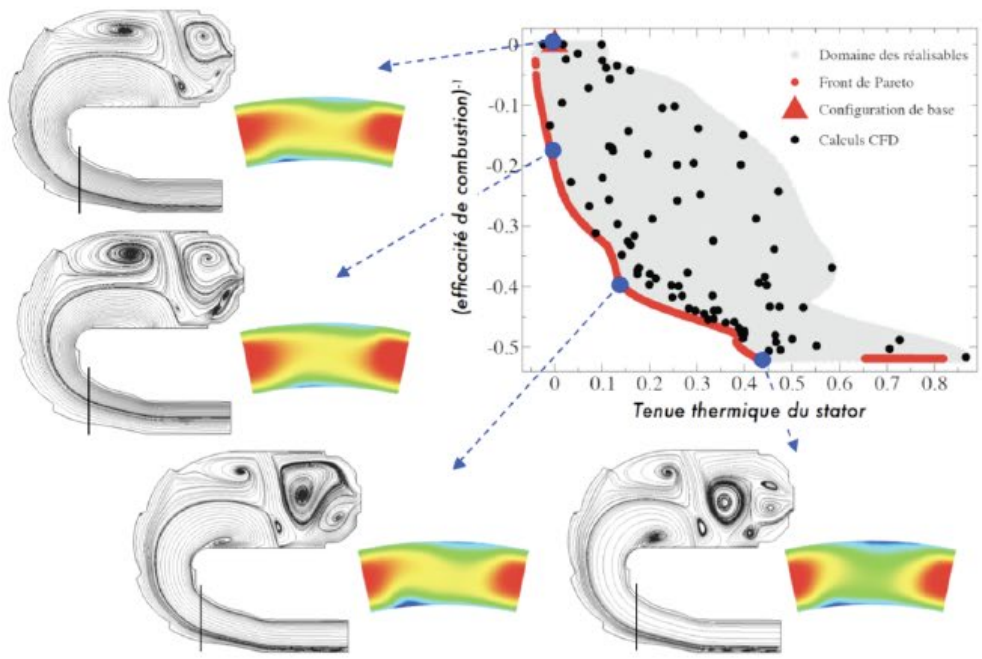
ONERA THE FRENCH AEROSPACE LAB

Code coupling

Why coupling codes?

➔ To construct applications around existing codes
- •On the fly post-processing,
- •Optimization loop around a code,
- •Ensemble simulations,
- •...





OpenPALM s'impose peu à peu pour le couplage de codes dans les contextes multiphysiques / multicomposants parallèles.

Pour y parvenir, Legolas++ se livre à une double opération : d'une part la composition verticale qui accole des matrices et d'autre part la composition horizontale qui les additionne. Au final, on peut reconstruire des matrices assez sophistiquées à partir d'un nombre fini de blocs, et résoudre le tout avec un solveur parallèle. La performance obtenue est très intéressante. En utilisant les bons outils, Legolas++ permet d'obtenir un *speedup* de 10X sur une machine multicœur à mémoire partagée.

Il faut également citer **FreeFem++**, un logiciel libre qui permet de résoudre les équations aux dérivées partielles. Créé - entre autres - par Frédéric Hecht (UPMC), il est, comme son nom l'indique, gratuit et basé sur une Finite Element Method. Au contraire de Legolas++ évoqué précédemment, le "++" ne fait pas ici référence au C++, mais au fait qu'il s'agit de sa troisième version.

Multiplateforme (il est disponible sous Windows, Mac et Unix), ff++ fonctionne par méthode d'éléments finis R2 et R3, avec un langage dédié, un seul paradigme de parallélisation (MPI, en l'occurrence) et ne prend en charge que les triangles et les simplexes, afin de conserver une structure "simple". Il gère des valeurs réelles ou complexes, récupère toutes les matrices et fait... tout ce que l'on peut attendre d'un logiciel d'éléments finis, à quoi il ajoute l'interpolation automatique des données sur des maillages distincts.

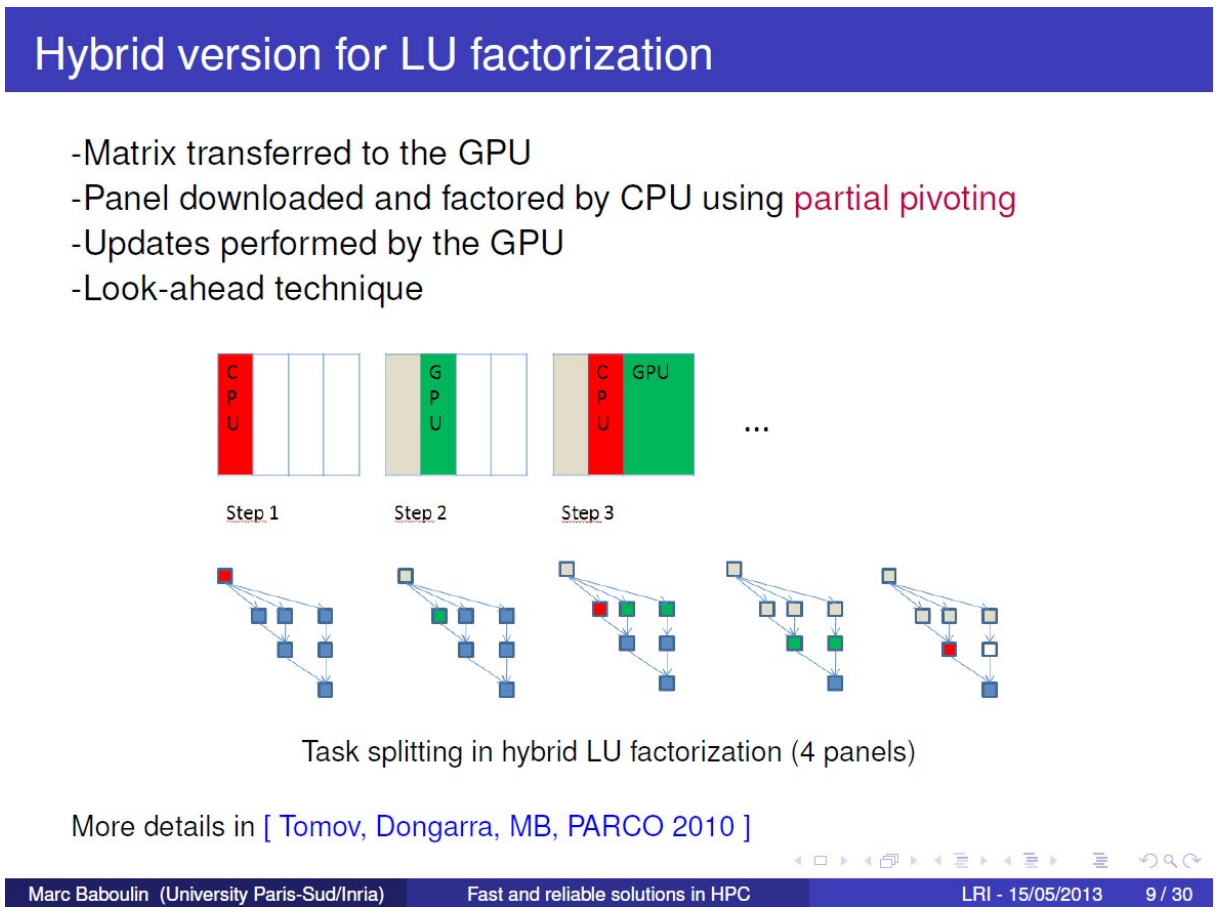
N'oublions pas non plus **OpenPALM**, coupleur co-developpé par le Cerfacs et l'ONERA. Florent Duchaine (Cerfacs) recommande ses performances, notamment en mécanique des fluides, pour réaliser des couplages multiphysiques, et en algorithmique dynamique. Son principe est d'utiliser du couplage de code pour traiter un système dans sa globalité avec d'un côté les aspects couplages physiques, et de l'autre les aspects informatiques avec des échanges de données. Un étage de turbine, par exemple, comportant une partie fixe et une partie mobile couplées, peut être traité par une instance du même code. OpenPALM s'utilise également dans le domaine du réchauffement climatique ou le couplage thermique, notamment pour le spatial.

Quelles que soient les bibliothèques, il est intéressant de considérer l'usage qu'en font les utilisateurs compte tenu des contraintes auxquelles ils sont soumis. Au CEA par exemple, les échelles de temps sont très longues, de sorte que les codes associés doivent avoir une durée de vie extensive. Les réacteurs que le CEA exploite (propulsions navales, systèmes nucléaires électrogènes et expérimentaux...) ont des cycles de vie qui s'étalent sur 20 ou 30 ans et obligent à couvrir un spectre très large de sciences. Voilà pourquoi, comme le précise Christophe Calvin, le CEA utilise principalement la plateforme **Salome**, un coupleur d'applications développé en partenariat avec EDF.

Quelles évolutions à terme ?

Constatant que trop d'énergie était dissipée sous forme de chaleur, les constructeurs ont arrêté la course au GigaHertz au tournant du siècle dernier. La bataille se joue désormais autour des processeurs multicœurs et des accélérateurs, notamment GPU. Mais pour Marc Baboulin, (Université Paris-Sud, INRIA), ces progrès engendrent une augmentation de l'hétérogénéité des architectures HPC, ce qui rend les calculateurs plus difficiles à utiliser.

C'est dans ce contexte que s'inscrit **MAGMA**, bibliothèque numérique qui vise à conserver l'interface habituelle de Lapack tout en fournissant des algorithmes efficaces sur les nouvelles architectures. MAGMA consiste à placer la matrice sur le GPU pour réduire le goulot d'étranglement des communications GPU/CPU. Il s'agit d'un parallélisme de niveau BLAS qui remplace les appels aux BLAS par des appels à CUBLAS tout en gardant une conception Lapack. Cette technique permet de garder sur le CPU les tâches qui ne tirent pas avantage du GPU, donc d'utiliser l'un et l'autre en fonction de leur meilleur potentiel respectif. L'asynchronisme CPU-GPU est alors une nécessité.



La bibliothèque MAGMA exploite les GPU pour optimiser les calculs matriciels complexes via CUDA.

Marc Baboulin n'est pas le seul à travailler sur la réduction des communications. Laura Grigori (INRIA) et son équipe en ont également fait une spécialité. Leur approche se fonde sur une nouvelle classe d'algorithmes d'algèbre linéaire issue de progrès théoriques récents dans le domaine. Ces nouveaux algorithmes *Communication Avoiding* permettent de minimiser le volume des communications (pour la bande passante) et le nombre de messages (pour la latence) mais ils nécessitent souvent d'effectuer des calculs supplémentaires, calculs toutefois largement compensés par les gains de temps. Laura Grigori précise que leurs modèles ne prennent pas (encore) en compte la distance entre les processeurs, élément qui peut impacter la durée des communications, mais qui augmenterait plus encore la complexité de l'optimisation.

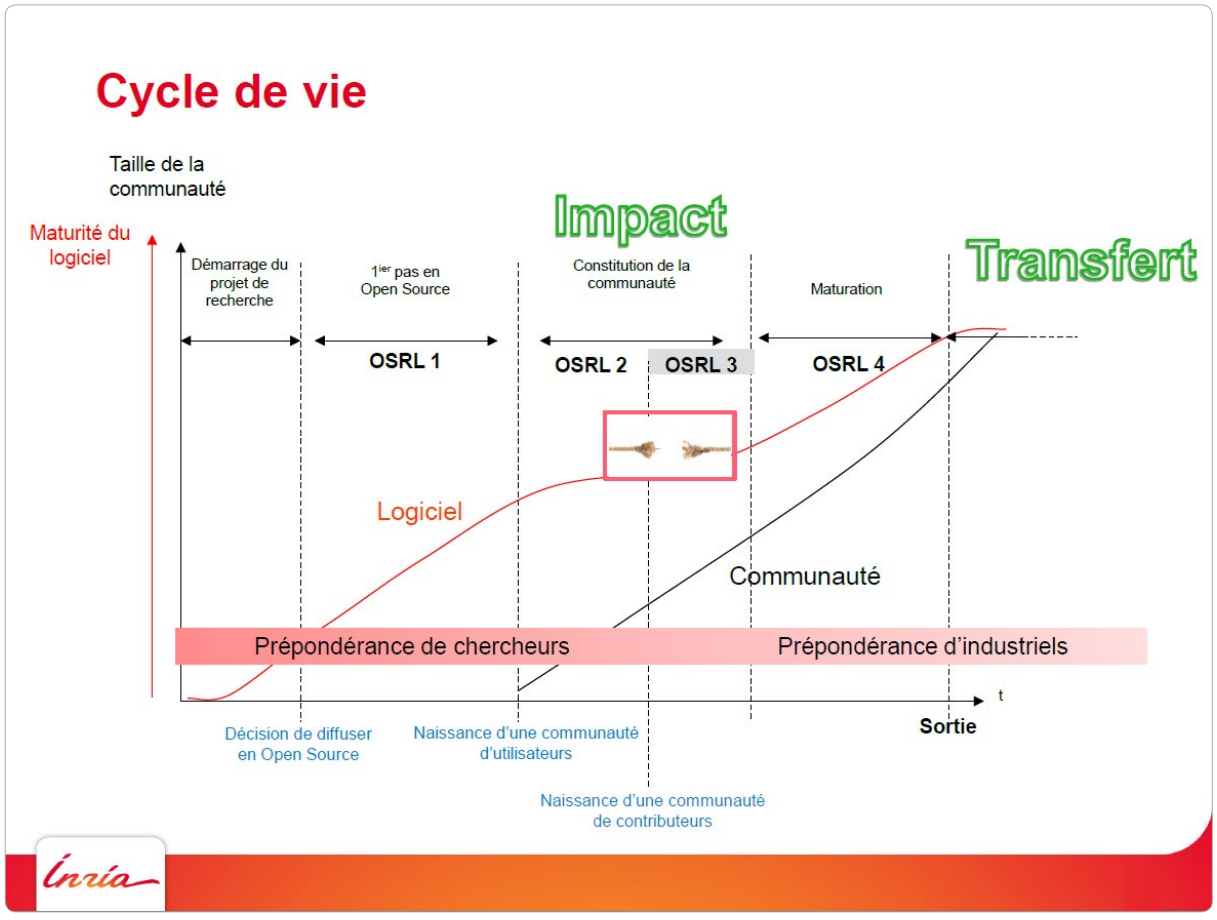
Les volumes de données à manipuler et à calculer en contexte HPC posent enfin de véritables défis en matière d'estimation et de contrôle des erreurs d'arrondi. C'est ce à quoi travaillent Fabienne Jézéquel et son équipe (LIP6-UPMC), autour de

la bibliothèque CADNA. Fondée sur la méthode CESTAC, CADNA exécute le même calcul trois fois avec différentes propagations d'erreurs d'arrondi : les chiffres qui diffèrent à l'issu des opérations sont considérés comme non significatifs. LIP6 travaille d'une part à gérer en parallèle ces trois exécutions sur les processeurs multicœurs, et d'autre part à permettre la validation numérique de codes parallèles utilisant MPI, CUDA ou encore OpenMP.

Le développement de bibliothèques en pratique

Le travail que représente le développement d'une bibliothèque logicielle, notamment pour le calcul, présente un certain nombre de spécificités. La dimension haute performance du HPC nécessite d'abord que l'outil en lui-même ne génère pas de pertes - ce qui ne va pas de soi. Par ailleurs, l'intérêt général est de construire des bibliothèques sémantiquement les plus homogènes possibles, qui elles-mêmes s'appuient sur d'autres bibliothèques de façon à ce que les progrès se propagent naturellement.

En termes de reconnaissance par les pairs - question importante compte tenu de l'investissement de temps - le développement d'un outil logiciel reste un excellent moyen de se faire connaître. FreeFem, par exemple, a au départ été développé pour des besoins de recherche internes. Aujourd'hui, il est utilisé par la communauté. Au final, la reconnaissance



Le concept d'Open Source Readiness Levels (OSRL) formalisé par Patrick Moreau (INRIA) permet d'appréhender puis d'anticiper les grandes étapes du cycle de vie d'une bibliothèque logicielle.

passe nécessairement par les publications, notamment pour apporter la preuve mathématique de faisabilité du code fonctionnel.

Cela étant, le fait que le code source ne soit généralement pas joint aux publications qui décrivent les résultats et les performances pose problème dans la mesure où il est impossible de refaire l'expérience et encore moins de rattraper les vingt années de développement pour les vérifier ! Le principe de réfutabilité de Popper est *de facto* violé. Il est donc nécessaire de mettre en place des démarches dans ce sens, avec des licences certes restrictives, mais qui respectent la démarche scientifique.

En ce qui concerne la durabilité d'un projet, Patrick Moreau (INRIA) propose le concept d'*Open Source Readiness Levels* (OSRL), sorte de ligne de vie d'un logiciel. Il distingue 4 phases :

• à OSRL 1, seuls les collègues du créateur utilisent le logiciel ;

• à OSRL 2, une communauté d'utilisateurs se forme ;

• à OSRL 3, phase optionnelle, les mêmes utilisateurs qu'à ORSL 2 ont la possibilité de contribuer au logiciel ;

• à OSRL 4, les industriels se joignent au cortège, avec leurs exigences, notamment en termes de corrections de bogues !

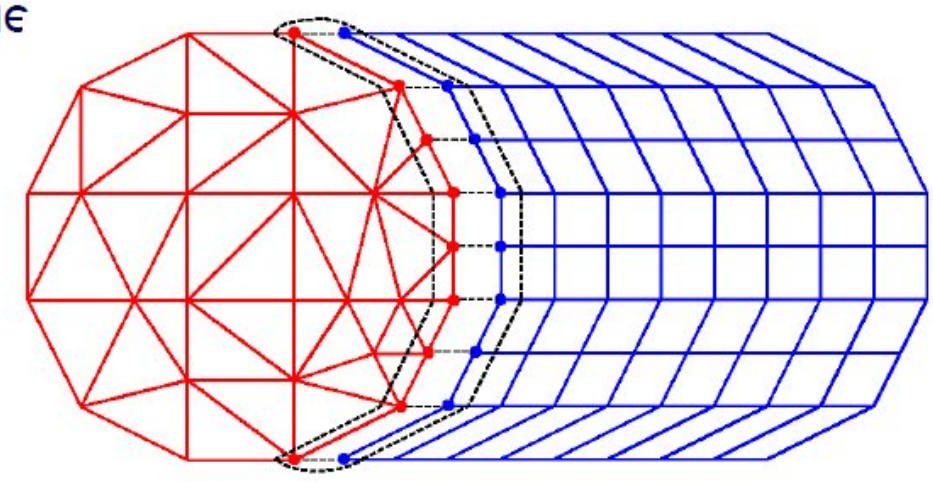
La phase OSRL 4 est censée aboutir au transfert technologique, qui pour être géré correctement conduit à l'idée de création d'un consortium. Patrick Moreau conseille de faire simple, via une association type Loi de 1901 et un simple contrat

Equations locales, conditions de raccord

- Equations internes dans chaque sous-domaine

$$K_{11}\, x_1 + K_{13}\, x_3^{(1)} = b_1$$

$$K_{22}\, x_2 + K_{23}\, x_3^{(2)} = b_2$$



- Condition d'admissibilité sur l'interface

$$x_3^{(1)} = x_3^{(2)} \quad (= x_3)$$

- Condition d'équilibre à l'interface

$$K_{31}\, x_1 + K_{32}\, x_2 + K_{33}\, x_3 = b_3$$

$$K_{31}\, x_1 + K_{33}^{(1)}\, x_3^{(1)} - b_3^{(1)} + K_{32}\, x_2 + K_{33}^{(2)}\, x_3^{(2)} - b_3^{(2)} = 0$$

FETI permet de décomposer un système complexe en sous-domaines, après l'avoir maillé par exemple avec Scotch.

d'adhésion. Il préconise également de protéger les marques et logos car ces éléments contribuent généralement à la valeur du consortium.

La délicate question des licences

Les fondements du mouvement du logiciel libre sont les mêmes que ceux qui président au développement scientifique, à savoir permettre le partage des idées et découvertes dans le but de faire progresser la recherche et le savoir collectif. C'est pour cette raison que le terme "logiciel libre" (en anglais *free software*) désigne tout logiciel garantissant à l'utilisateur les quatre libertés fondamentales suivantes :

• pouvoir exécuter (utiliser) le programme, pour n'importe quel usage ;

• pouvoir étudier le fonctionnement du programme (ce qui nécessite l'accès à son code source) et l'adapter à ses propres besoins ;

• pouvoir redistribuer des copies ;

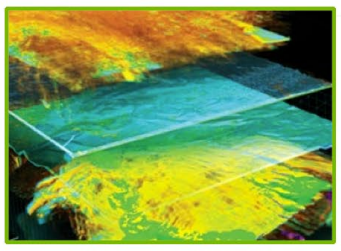
• pouvoir améliorer le programme et diffuser ces améliorations pour en faire profiter la communauté.

Il ne faut donc pas confondre logiciel libre avec logiciel gratuit (*freeware*) ni avec le concept de logiciel Open source. Le terme "Open source" désigne tout logiciel dont le code source est accessible à l'utilisateur, mais cela n'en fait pas pour autant un logiciel "libre". Si tous les logiciels libres sont par définition Open source, tous les logiciels Open source ne sont pas forcément libres, c'est-à-dire qu'ils ne garantissent pas nécessairement l'ensemble des quatre libertés citées précédemment. Pour lever toute ambiguïté, les anglophones utilisent parfois le terme "free/libre open-source software" (FLOSS) pour désigner les logiciels à proprement parler libres. Dans tous les cas, comme le souligne Patrick Moreau, il est recommandé de se déterminer sur telle ou telle licence pour son code Open Source dès qu'on en sait assez sur la nature et la portée du projet. ■

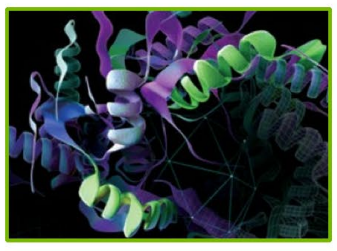
CAPS

The Many-Core Programming Company

x5,6 en 5 jours
Comprendre les origines de l'univers
Panos Labropoulos, Université de Groningen et ASTRON

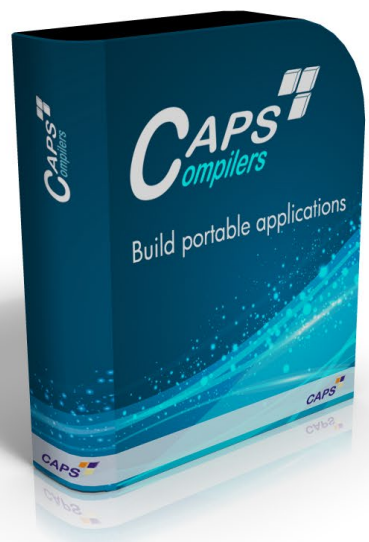


x3 en 7 jours
Trouver du pétrole dans les plus grands réservoirs du monde
Global Oil Company



x5 en 1 jour
Simuler des solvants dans les protéines
Bharat Medasani, Université du Texas à San Antonio

NVIDIA®



Simple.

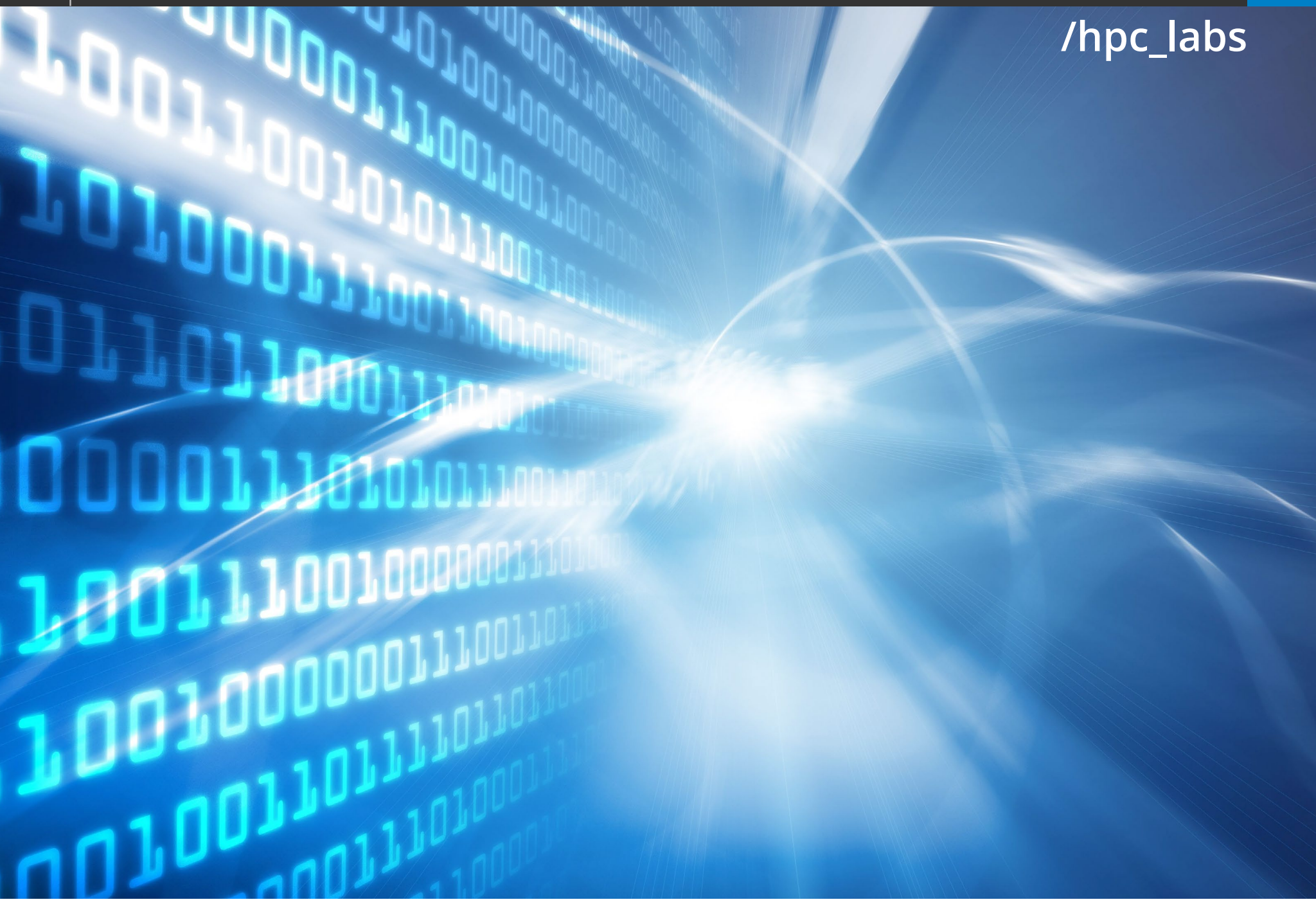
Puissant.

Ouvert.

Profitez d'une évaluation gratuite de 30 jours de

CAPS OpenACC Compilers

En savoir plus *



/hpc_labs

MAITRISER OPENACC 2.0 [1ÈRE PARTIE]

LES NOUVELLES FONCTIONS DE GESTION DE DONNEES

Après une première release qui a beaucoup fait pour faciliter le portage du calcul parallèle sur les accélérateurs - et pour le standardiser - OpenACC connaît aujourd'hui une mise à jour majeure. Penchons-nous donc sans plus attendre sur ses améliorations très attendues, en commençant cette série de trois articles par les nouvelles fonctions de gestion de données - un domaine qui méritait vraiment d'être repensé...

STEPHANE CHAUVEAU, PHD

C'est en 2011 que PGI, Cray et NVIDIA ont, avec l'aide de CAPS, sorti la première spécification d'OpenACC. Il s'agissait alors d'une première mouture de la future extension de directives pour accélérateurs d'OpenMP. Après deux ans d'améliorations continues apportées notamment par les nouveaux entrants au consortium, la version 2.0 d'OpenACC vise à devenir le standard de fait de la programmation des accélérateurs par directives en C/C++ et Fortran.

L'une des faiblesses les plus régulièrement reprochées à OpenACC 1.0 est son manque de flexibilité dans la gestion des données. Dans cette version initiale, le chargement de données sur l'accélérateur s'effectue en créant une région devant être parfaitement imbriquée dans le code. Le constructeur **data** est le moyen le plus simple de créer de telles régions, mais d'autres sont aussi implicitement créées autour des constructeurs de

calcul (**parallel** et **kernels**). On peut également passer par la directive indépendante **declare** pour définir une région dans le contexte où la directive est utilisée - région qui peut être soit locale à la procédure courante, soit globale à l'application entière. Dans tous les cas, en OpenACC 1.0, les règles de chargement des variables suivent celles du contexte des variables du langage source.

Ce choix ne manque pas de sens du point de vue du compilateur sachant que, par construction, la durée de vie d'une région sur l'accélérateur ne peut dépasser la durée de vie de la variable à laquelle elle réfère sur le processeur hôte. Mais une telle stratégie pose des problèmes dans des applications qui utilisent des structures de données plus complexes. OpenACC 1.0 ne fournit par exemple aucun moyen de gérer un nombre dynamique de données sur l'accélérateur - à part bien sûr quelques astuces telles que la création d'une fonction récursive pour encapsuler une région de données autant de fois que nécessaire, ce qui manque clairement d'élégance. Il est par ailleurs impossible d'écrire du code dans lequel les données sur l'accélérateur sont allouées et libérées depuis des fonctions différentes. Or, c'est typiquement le cas des applications orientées objets où ces manipulations sont logiquement réalisées dans les constructeurs et destructeurs.

Listing 1 : Traitement inefficace de paires avec OpenACC 1.0.

```
void doit( size_t n, struct Value * data[n] )
{
 for (size_t i=0;i<n;++i)
 {
  struct Value * data1 = data[i] ;
  #pragma acc data copy( data1[0:1] )
  {
   for (size_t j=i+1;j<n;++j)
   {
    struct Value * data2 = data[i] ;
    #pragma acc parallel copy( data2[0:1] )
    {
     work_on_a_pair( data1 , data2 ) ;
    }
   }
  }
 }
}
```

Listing 2 : Chargement dynamique avec OpenACC 2.0.

```
void doit( size_t n, struct Value * data[n] )
{
 for (size_t i=0;i<n;++i)
 {
  struct Value * data = data[i] ;
  #pragma acc enter data copyin(data[0:1])
 }

 for (size_t i=0;i<n;++i)
 {
  for (size_t j=i+1;j<n;++j)
  {
   struct Value * data1 = data[i] ;
   struct Value * data2 = data[j] ;
   #pragma acc parallel present( data1[0:1], data2[0:1] )
   {
    work_on_a_pair( data1 , data2 ) ;
   }
  }
 }

 for (size_t i=0;i<n;++i)
 {
  Value * data = data[i] ;
  #pragma acc enter data copyout(data[0:1])
 }
}
```

Ce dernier cas est illustré au **listing 1.** Une fonction doit traiter par paires un nombre arbitraire **n** de structures de données fournies par un vecteur de pointeurs. Etant donné qu'OpenACC 1.0. ne fournit aucune solution pour charger les **n** structures simultanément, le

code ne garde en mémoire que deux structures à la fois, ce qui n'est évidemment pas satisfaisant puisque chacune des **n** structures doit être chargée en moyenne **n**/2 fois. Les performances pourraient sans doute être améliorées en utilisant l'astuce récursive mentionnée plus haut, ou en gérant plus d'une structure par région de données, mais cela nécessiterait de modifier en profondeur le code d'origine, ce qui n'est pas exactement l'objectif d'une API de programmation par directives telle qu'OpenACC.

Listing 3 : Implémentation personnalisée de present_or_copy(A[n]) utilisant l'API.

```
int A[n] ;

...

/* A[0:n] may or may not be present on the device */
int was_present ;
was_present = acc_is_present( A , n*sizeof(int) );
if ( !was_present ) {
  acc_copyin( A , n*sizeof(int) )
}

...

 /* A[0:n] is now present on the device */

 ...

 if ( !was_present ) {
   acc_delete( A , n*sizeof(int) )
 }

 /* A[0:n] may or may not be present on the device */

 ...
```

La gestion de données devient dynamique

OpenACC 2.0 résout la plupart de ces problèmes en introduisant de nouvelles fonctionnalités de gestion dynamique des données. Deux directives indépendantes, **enter data** et **exit data**, peuvent être appelées pour respectivement créer ou détruire les données sur l'accélérateur. Utilisées ensemble, elles correspondent à une construction **data**, à ceci près qu'elles ne nécessitent aucune imbrication. Elles peuvent donc être placées n'importe où dans le code et, contrairement aux autres constructeurs, peuvent même être exécutées de façon asynchrone.

Le **listing 2** montre comment le code de notre premier exemple peut être nettement optimisé avec OpenACC 2.0. Chacune des **n** structures est maintenant chargée sur l'accélérateur en une seule fois (un **copyin** et un **copyout**). Notez que les directives **enter data** et **exit data** ont leur équivalent dans l'API (appels **acc_copyin**, **acc_copyout**, **acc_delete**, etc.).

Cela étant, comme du point de vue langage les directives **enter data** et **exit data** ne créent pas une région de données en tant que telle, la clause **present** reste nécessaire dans le constructeur **parallel** pour rendre les données accessibles depuis la région parallèle. Cette clause ne provoque aucune affectation ni aucun transfert de données. Les directives **enter data** et **exit data** ainsi que leurs équivalents dans l'API sont donc particulièrement adaptées à l'écriture de bibliothèques dans lesquelles les données sont gérées et utilisées par de multiples fonctions. La plupart des fonctionnalités de gestion de données présentes dans les directives ou les constructeurs sont maintenant également proposées dans l'API. C'est le cas par exemple de **acc_copyin()**, **acc_present_or_copyin()**, **acc_create()** et **acc_present_or_create(),** qui correspondent aux clauses autorisées dans la directive **enter data**, tandis qu'**acc_copyout()** et **acc_delete()** correspondent aux clauses de la directive **exit data**. Toutes ces fonctions sont malheureusement synchrones mais il est probable que des alternatives asynchrones seront disponibles dans une prochaine version d'OpenACC.

Les développeurs qui souhaiteraient implémenter leurs propres fonctions de gestion des données trouveront ce qu'il

faut parmi les fonctions de bas niveau de l'API : **acc_malloc()** et **acc_free()** allouent directement la mémoire sur l'accélérateur, **acc_map_data()** et **acc_unmap_data()** associent les adresses des variables sur le processeur et sur l'accélérateur, **acc_deviceptr()** résout l'adresse accélérateur associée à une adresse CPU cependant que **acc_hostptr()** réalise l'opération l'inverse. Enfin, **acc_memcpy_to_device()** et **acc_memcpy_from_device()** transfèrent des régions mémoire arbitraires entre hôte et accélérateur.

Les directives de données indépendantes (et leurs appels API correspondants) peuvent être utilisées avec les constructeurs de données mais les données présentes sur l'accélérateur qui ont été créées avec un certain type de directive ne peuvent être détruites avec un autre. Les directives de gestion des données n'ont pas connaissance les unes des autres (souvenez-vous, elles sont "indépendantes"), de sorte qu'il serait incorrect de vouloir changer la sémantique d'une clause **present_or_copyin** simplement en appelant **acc_present_or_copyin()** puis **acc_delete()**.

La bonne façon de faire, illustrée au **listing 3**, implique de ne réaliser ces deux appels que si les données ne sont pas déjà présentes sur l'accélérateur. Cette restriction est à la fois saine et logique. Elle devrait éviter en pratique un certain nombre d'erreurs voire d'effets de bord qui ne se manifesteraient qu'à l'exécution.

Listing 4 : Chargement d'un tableau multidimensionnel.

```
void doit( size_t nb, struct Value * data[nb] )
{
 #pragma acc parallel copy( data[0:nb][0:1] )
 {
  for (size_t i=0;i<nb;++i)
  {
   for (size_t j=i+1;j<nb;++j)
   {
    work_on_a_pair( data[j], data[i] ) ;
   }
  }
 }
}
```

Listing 5 : Variable globale à durée de vie dynamique sur l'accélérateur.

```
//---- external.c ----

int counter ;

#pragma acc routine(update_counter)
void update_counter(void) {
 counter++ ;
}


//---- main.c ----

extern int counter ;

void work(void)
{
 #pragma acc parallel
 {
  /*
   * Problem: The compiler is not aware that update_counter is
   * accessing counter. Since counter has a dynamic lifetime on
   * the device, how can update_counter figure out its (non-constant)
   * device address?
   */
  update_counter() ;
 }
}

int main(void)
{
 #pragma data copy(counter) ;
 {
  work() ;
 }
}
```

Implémenter des tableaux multidimensionnels en tant que vecteurs de pointeurs

Sachant qu'implémenter des tableaux multidimensionnels avec des vecteurs de pointeurs est assez courant en C/C++, la pratique est maintenant directement supportée par OpenACC 2.0, au moins pour les tableaux réguliers. Avec une seule directive pour spécifier la forme du tableau multidimensionnel, le compilateur prendra en charge toute la mécanique de création d'un tableau de pointeurs équivalent sur l'accélérateur. Il existe toutefois une restriction importante : les pointeurs formant le tableau doivent tous être soit présents soit non présents. Autrement dit, modifier les pointeurs sur l'hôte ou sur l'accélérateur n'est pas autorisé.

Cette fonctionnalité peut être appliquée directement à notre premier exemple dans lequel le tableau de pointeurs **data[n]** sera traité par OpenACC 2.0 comme un tableau de structures bidimensionnel **data[n][1]**. C'est ce que montre le **listing 4**, où le gain en termes de productivité et de lisibilité par rapport au **listing 2** apparaît évident.

La simplicité de cette syntaxe ne doit cependant pas faire oublier que la plupart des implémentations généreront un transfert de données par pointeur. Ainsi, les performances sur un tableau non contigu seront significativement moins élevées qu'avec un tableau contigu. Le code pourra cependant bénéficier d'un effet de bord intéressant pour la gestion du tableau en tant que tel : les boucles utilisées pour itérer sur les pointeurs pouvant maintenant être exécutées directement sur l'accélérateur, il en résultera quelquefois une meilleure mise à profit de son parallélisme matériel.

Listing 6 : Bibliothèque I/O utilisant la nouvelle fonctionnalité link.

```
//---- io.c ----

#pragma acc declare copyin(io_status, io_size)
#pragma acc declare link(io_buffer)
int  io_size   = 0 , io_status = FALSE ;
char io_buffer[10000000] ;

#pragma routine(io_putchar)
void io_putchar(char c)
{
  if ( io_status == TRUE ) {
    io_buffer[ io_size++ ] = c ;
  }
}

void io_start()
{
  io_status = TRUE ;
  io_size   = 0 ;
  #pragma acc update device(io_size, io_status)
  #pragma acc enter  data create(io_buffer)
}

void io_stop()
{
  io_status = FALSE ;
  #pragma acc update self(io_size) device(io_status)
  #pragma acc exit data copyout(io_buffer)
  fwrite(io_buffer, io_size, 1 , stdout) ;
}


//---- main.c ----

void main()
{
  io_start() ;
  #pragma acc parallel
  {
    ...
    io_putchar('X') ;
    io_putchar('Y') ;
    io_putchar('Z') ;
    ...
  }
  io_stop() ;
}
```

Le support d'autres types de données complexes tels que les listes chaînées, les arborescences ou les structures contenant des pointeurs est prévu dans les prochaines versions d'OpenACC. En attendant, il est déjà possible de gérer des données complexes en utilisant les nouveaux appels API présentés précédemment.

Les données globales

Enfin, l'un des apports les plus importants de cette nouvelle version est sans doute le support amélioré des variables globales par la directive **declare**. Avec OpenACC 1.0, il était déjà possible de créer une région de données globale en utilisant une des clauses standard (**create**, **copy**, **copyin**, par exemple), tandis que la clause **device_resident** servait à créer les variables sur l'accélérateur sans qu'elles existent sur le processeur hôte. Les variables globales ayant une durée de vie dynamique sur l'accélérateur étaient supposées être gérées comme toutes les autres variables, c'est-à-dire dans des régions de données explicites ou implicites bornant les directives de calcul.

En version 2.0, OpenACC supporte officiellement les appels aux fonctions utilisateurs dans les directives de calcul. C'est d'ailleurs un problème car lorsque le compilateur traite ces directives, il peut ne pas être en mesure de déterminer quelles variables globales à durée de vie dynamique seront utilisées par la fonction appelée. Cette situation est illustrée au **listing 5** où une fonction non locale, appelée depuis une région de calcul, utilise une variable globale dont la durée de vie est dynamique sur l'accélérateur.

En introduisant une nouvelle clause **link** dans la directive **declare**, OpenACC 2.0 apporte une solution simple à ce cas de figure. Une variable globale déclarée en tant que variable **link** sera toujours présente sur l'accélérateur sous la forme d'un lien global qui, initialement, ne pointe sur rien – un peu comme un pointeur initialisé à **NULL**. Ce lien sera mis à jour de façon dynamique pour pointer vers la bonne adresse ou vers **NULL** à chaque fois que le lien global sera utilisé par un constructeur **data** ou une directive **enter data** ou **exit data**.

Restriction importante sur les variables **link** globales, elles ne peuvent pas être utilisées avec certains appels API tels que **acc_create()** ou **acc_copy()**. Ces appels n'acceptant qu'une adresse hôte, ils ne peuvent pas connaître la nature **link** de la variable correspondante. De ce fait, logiquement, on ne peut pas mettre à jour l'adresse de la variable sur l'accélérateur.

Un des principaux inconvénients des variables **link** globales par rapport aux variables régulières est qu'il est possible, mais évidemment interdit, d'y accéder au niveau de l'accélérateur lorsqu'elles ne sont gérées par aucune région de données. De telles erreurs ne peuvent généralement pas être interceptées à la compilation et, dans la plupart des cas, elles aboutissent à une erreur d'exécution identique à ce qui surviendrait lors de l'utilisation d'un pointeur illégal sur l'accélérateur.

Pour illustrer l'utilisation des clauses **link**, le **listing 6** en présente un cas d'usage typique - une bibliothèque I/O très simplifiée dans laquelle l'énorme buffer de sortie est géré de façon dynamique en tant qu'objet **link**.

Vous devriez maintenant être en mesure de tirer profit des nouvelles fonctions de gestion de données d'OpenACC 2.0 - et de les implémenter dans vos codes assez simplement. Le mois prochain, nous nous focaliserons sur d'intéressantes améliorations des constructeurs de calcul : constructeurs atomiques, clauses spécifiques à certains accélérateurs, parallélisme imbriqué, appels de routines, etc. Restez à l'écoute et, en attendant...

Bons développements !

Abonnez-vous*

[gratuitement]

Recevez chaque mois le seul magazine exclusivement dédié aux environnements HPC et Big Data...

...et profitez des contenus Web qui vous sont réservés

*sur **www.hpcmagazine.fr**