

SUPERCALCULATEURS

Aux extrêmes du calcul

Les nouveaux horizons

Les grands challenges

Futur : vers l'Exascale



ÉDITORIAL

3

SUPERCALCULATEURS

L'ENJEU DU FUTUR

Le calcul scientifique de haute performance, ou HPC, s'est peu à peu installé dans notre vie, sans que nous en ayons toujours conscience. Il est dans nos médicaments, nos placements financiers, les films que nous voyons au cinéma et l'équipement de nos champions, la voiture que nous conduisons et l'essence qui la fait fonctionner. Il rend notre monde plus sûr, plus économe de ses ressources et, grâce à nos chercheurs, plus compréhensible.

Mais ces pas de géants que nous avons accomplis, notamment en franchissant la barre du pétaflops, vont nous paraître vite bien modestes, car les plus extraordinaires bouleversements

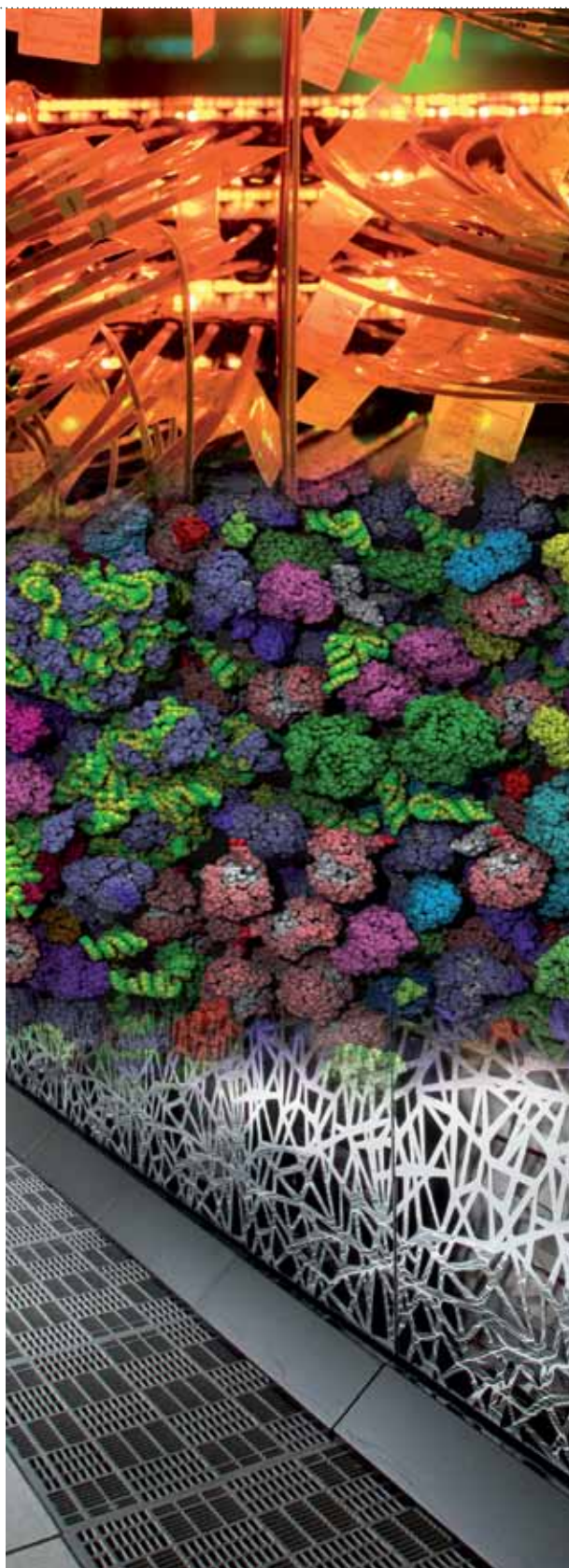
technologiques sont devant nous : le Cloud computing, qui va révolutionner et démocratiser l'accès à la très haute puissance, l'exaflops, qui fera entrer la simulation numérique dans une nouvelle dimension...

Pour la recherche, c'est la promesse de découvertes inimaginables. Pour les entreprises, de toujours plus d'innovation, d'optimisation et de sécurité. Pour les États, enfin, ce pourrait être le facteur clé de la compétitivité et de l'indépendance dans le jeu international. Or, développer les systèmes et les applications de la génération exaflopique nécessitera la mobilisation et la coopération de toute la communauté industrielle, scientifique et universitaire.

Demain, la recherche, l'innovation, l'industrie et l'emploi auront plus que jamais besoin de HPC. Pour l'Europe, c'est l'enjeu du futur.

PHILIPPE
VANNIER

Président-
directeur
général de Bull



SOMMAIRE

LA Recherche

74, AVENUE DU MAINE, 75014 PARIS
SOPHIA PUBLICATIONS (L'HISTOIRE)

PRÉSIDENT-DIRECTEUR GÉNÉRAL
Philippe Clerget

DIRECTRICE DE LA RÉDACTION
Valérie Hannin

CONSEILLERS DE LA DIRECTION
Michel Winock
Jean-Noël Jeanneney

RÉDACTRICE EN CHEF
Séverine Nikel
(responsable des hors-séries)

RÉDACTRICE
Géraldine Soudri

A COLLABORÉ À CE NUMÉRO
Légendes Cartographie

**DIRECTION ARTISTIQUE ET
RÉALISATION À NOIR**
Correctrice Véronique Duthille

ICONOGRAPHIE
Jérémy Suarez

RESPONSABLE DE FABRICATION
Christophe Perrusson

**DIRECTRICE DU MARKETING DIRECT
ET DES ABONNEMENTS**
Virginie Marliac

CHARGÉE DU MARKETING DIRECT
Estelle Castillo

DIRECTRICE VENTES AU NUMÉRO
Évelyne Miont

**DIRECTRICE COMMERCIALE PUBLICITÉ
ET DÉVELOPPEMENT**
Caroline Nourry

PUBLICITÉ SECTEUR CULTUREL
DIRECTRICE DE CLIENTÈLE
Françoise Hullot

Publicité secteur littéraire Directrice
de clientèle Marie Amiel

© 2011
Sophia Publications

LES NOUVEAUX HORIZONS

04 LE DÉFI PERMANENT
DES SUPERCALCULATEURS

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

08 DÉMOCRATISONS
LE CALCUL INTENSIF!

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

11 POLITIQUE INRIA

Ena, vite tem oc tudempoNem
crimus turo morum cussil

12 LE TERA 100 BRILLE CÔTÉ
RENDEMENT

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod
crimus turo morum cussil

14 LE FUTUR
DES PROCESSEURS

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

LES GRANDS CHANGEMENTS

16 MODÉLISER LES
MOLÉCULES DU VIVANT
POUR MIEUX SOIGNER

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

20 LE SUPERCALCULATEUR
AU SERVICE DE L'ALERTE
TSUNAMI

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

22 LES FUTURES
RÉACTEURS NUCLÉAIRES
PROFITENT DÉJÀ
DU CALCUL HPC

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

24 VOIR LES MATÉRIAUX
GRANDIR ATOME PAR ATOME

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod
crimus turo morum cussil

26 LE BON CALCUL DE LA
DISSUASION NUCLÉAIRE

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

28 COMPRENDRE COMMENT
SE FORMENT LES ÉTOILES

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

30 PHYSIQUE DES CHOCs

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

32 LES DÉFORMATIONS
MARTENSITIQUES SOUS LA
LOUPE DES PROCESSEURS

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

34 DES PROCESSEURS
GRAPHIQUES POUR
VISUALISER LA LUMIÈRE

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod

FUTUR: VERS L'EXASCALE

35 LE PROCHAIN DÉFI :
LA MAÎTRISE DE L'ÉNERGIE

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod
crimus turo morum cussil

41 LA QUESTION
DE L'ÉNERGIE

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod
crimus turo morum cussil

42 LA QUESTION
DE L'ÉNERGIE

Ena, vite tem oc tudempoNem
proreo, nonostio uritiam pultod
crimus turo morum cussil



1996 : LA FRANCE SIGNE L'ARRÊT DÉFINITIF DES ESSAIS NUCLÉAIRES. LE CEA DOIT ALORS RELEVÉ UN NOUVEAU CHALLENGE : GARANTIR, D'ICI À 2011, LA FIABILITÉ ET LA SÛRETÉ DES ARMES NUCLÉAIRES PAR LA SEULE SIMULATION. RETOUR SUR CES QUINZE ANNÉES D'AVENTURE INDUSTRIELLE ET DE RECHERCHE AVEC JEAN GONNORD, DU CEA.

« LE DÉFI PERMANENT DES SUPERCALCULATEURS »

● **Nous sommes en 2011. Avez-vous atteint vos objectifs ?**

Jean Gonnord : Nous venons de livrer aux concepteurs d'armes le « Standard 2010 » : l'ensemble des codes de simulation d'armes nucléaires qui, associé à notre supercalculateur, le Tera 100, désormais opérationnel, permettra de garantir les futures têtes nucléaires des sous-marins, sans nouvel essai nucléaire. Une première scientifique ! Seuls les États-Unis ont affronté, comme la France, cette problématique très ambitieuse.

● **Votre vision du calcul haute performance et de la simulation est aujourd'hui**



JEAN GONNORD

est le chef du projet simulation numérique et informatique à la Direction des applications militaires du CEA.

unanimentement partagée par l'industrie et la recherche...

J. G. : Heureusement ! L'Europe était très en retard. Entre 1996 et 2006, sa part dans le Top 500 était passée de 28 % à moins de 17 % pour revenir aujourd'hui à 25 %. Il nous a fallu dix ans pour convaincre que la simulation haute performance (HPC, pour High Performance Computing) était stratégique, tant pour le monde industriel – afin de réduire les cycles de développement et les coûts – que pour celui de la recherche – énergie, climat, santé, etc. C'est désormais une évidence dans le monde. Mais disposer d'une capacité de calcul ne suffit

pas. Si l'on considère cette capacité comme stratégique, il faut en contrôler la technologie ! Or, jusqu'en 2005, l'Europe était absente du marché du HPC : moins de 0,2 % des machines y était conçu. Notre vision pour cette industrie est globale : de la maîtrise des technologies matérielles et logicielles, de leur intégration dans des supercalculateurs, à l'application finale. Sur ce point, nous avons souvent, l'impression de prêcher dans le désert.

● **Comment avez-vous mis en œuvre cette politique ?**

J. G. : En ingénieur, avec une stratégie à long terme et des

Plus de 200 personnes au CEA et chez Bull ont travaillé ensemble pour concevoir et réaliser le Tera 100. C'est la troisième machine installée au CEA de Bruyères-le-Châtel (Essonne), dédiée à la simulation. Ci-dessous, des écoulements tourbillonnaires.

étapes intermédiaires pour profiter du retour d'expérience. Avec pour principe de développer des machines généralistes, fiables, qui supportent la concurrence, pas des bêtes de calcul pour faire la course à la première place, ni des machines d'arsenal pour notre seul besoin. Le très haut niveau d'expertise de l'équipe et la capacité du Commissariat à l'énergie atomique et aux énergies alternatives (CEA) à organiser de grands projets ont fait le reste. Tenir nos engagements en 2010 supposait de disposer de 500 teraflops^(*). En 2001, nous avons visé 5 teraflops et conçu le supercalculateur Tera 1 et, en 2005, 50 teraflops avec le Tera 10. Le succès du Tera 10 nous a poussés à doubler la puissance prévue pour le Tera 100, qui dispose finalement de plus de 1 000 teraflops – soit 1 pétaflops. Face à de telles

avons mis en place une politique très innovante : le codesign, qui associe le constructeur et l'utilisateur expert en architecture. Grâce au codesign, l'utilité réelle de ces Formule 1 de l'informatique est garantie.

● Comment cela s'est-il organisé avec le constructeur ?

J. G. : Sur la base d'un contrat, résultat d'un appel d'offres lancé en 2008, où nous propositions une R&D partagée – donc un partage de la propriété intellectuelle –, la réalisation d'un démonstrateur et une option d'achat. Bull, constructeur français, qui avait déjà réalisé le Tera 10, l'a emporté. Plus de deux cents personnes du CEA et de chez Bull ont travaillé ensemble sur le Tera 100, tant sur l'architecture matérielle que sur les logiciels système. Cela a été un immense succès humain et organisationnel, matérialisé par

qualité. C'est une machine généraliste et non un outil de recherche comme les Blue Gene ou le Roadrunner d'IBM, à Los Alamos au Nouveau-Mexique, que nous avons finalement battus. C'est enfin un vrai succès commercial pour l'industriel français Bull. L'architecture du Tera 100 a reçu le prix de la « Meilleure architecture de l'année 2009 » par le magazine de référence américain, HPCwire⁽¹⁾. Bull a vendu ses ordinateurs en Europe mais aussi au Brésil. Notre équivalent anglais (AWE) lui a acheté deux machines de 150 teraflops ; Genci a commandé pour le programme européen Prace la machine Curie de plus de 1,5 pétaflops, qui sera opérationnelle au Très Grand Centre de calcul du CEA (TGCC) à Bruyères-le-Châtel cette année – ce sera la première machine pétaflopique à la disposition de l'ensemble de la recherche européenne. Le programme mondial pour la fusion contrôlée F4E a commandé en avril un ordinateur de 1,5 pétaflops qui sera installée à Rokasho, au Japon. Cela prouve que lorsqu'on a un objectif, avec la volonté de le réaliser et une organisation sans faille, rien n'est impossible.

● Comment voyez-vous l'avenir ?

J. G. : Il s'agit maintenant de pérenniser cette capacité pour l'Europe, que nous avons démontrée, à concevoir et réaliser ces très grands ordinateurs, stratégiques pour notre économie et notre société toute entière. À l'heure où le Japon et la Chine occupent les premières places, devant les États-Unis, l'Europe ne peut rester la seule région du Monde à laisser à d'autres le contrôle d'une technologie qui est l'une des clés de son avenir. C'est pourquoi, nous soutenons la création d'un ETP pour European Technology Platform, piloté par les industriels européens et s'appuyant sur les grands laboratoires de recherche. En ce qui nous concerne, la R&D pour les deux prochaines générations de machines du CEA/DAM : Tera 1000 et EXA1 est d'ores et déjà lancée. Et nous serons au rendez-vous de l'exaflops^(*) avant la fin de cette décennie. ●

PROPOS RECUEILLIS PAR ISABELLE BELLIN

⁽¹⁾ HPCwire, novembre 2009.

puissances, une architecture massivement parallèle s'est imposée. Ce qui impliquait pour le Tera 100 plus de 100 000 processeurs de haute technologie. Pour des raisons économiques, nous avons utilisé des composants fabriqués en très grandes quantités, dits « composants pris sur étagère ». Et, quinze ans plus tard, ils s'avèrent aussi être les plus performants. Seul le marché mondial permet de financer la Recherche et développement nécessaire à un nouveau processeur. Côté logiciels, nous avons choisi de mutualiser développement et validation en imposant des logiciels libres. Enfin, pour le Tera 100, nous

la création d'un laboratoire commun, Extreme Computing, implanté sur trois sites : aux Clayes-sous-Bois (Yvelines), à Échirolles (Isère) et à Bruyères-le-Châtel (Essonne), où est située Ter@tec, la première technopole européenne du HPC.

● Quel bilan tirez-vous de ces quinze années ?

J. G. : Le Tera 100 est la première machine conçue et réalisée en Europe à dépasser le pétaflops. L'ensemble du programme défini en 1996 a été réalisé dans les délais et les budgets annoncés. Avec un rendement^(*) exceptionnel de 83,7 %, l'un des meilleurs du Top 500, le Tera 100 démontre sa fiabilité et sa



LE FLOPS
(Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un teraflops permet de faire mille milliards d'opérations par seconde (10^{12}) et un exaflops un milliard de milliards par seconde (10^{18}).

Le rendement est le rapport entre la puissance mesurée et la puissance théorique.

LA SIMULATION NUMÉRIQUE EST PROGRESSIVEMENT DEVENUE UN VECTEUR UNIVERSEL DE DÉVELOPPEMENTS SCIENTIFIQUES ET ÉCONOMIQUES. L'AMBITION DE LA STRUCTURE PUBLIQUE GENCI EST DE DÉMOCRATISER L'UTILISATION DU CALCUL INTENSIF EN LE RENDANT ACCESSIBLE À TOUS LES SCIENTIFIQUES SUR LE TERRITOIRE NATIONAL.

DÉMOCRATISONS LE CALCUL INTENSIF!

Imaginez des machines si puissantes qu'elles effectuent en une seule journée une tâche qu'un ordinateur de bureau mettrait au moins 150 ans à accomplir... Science-fiction? Non, science tout court! Ces machines, que l'on appelle des supercalculateurs, sont capables d'effectuer des millions de milliards d'opérations en une seule seconde – d'où le terme de calcul intensif. Elles permettent de reproduire, par la modélisation et la simulation, des expériences qui ne peuvent pas être réalisées en laboratoire, quand elles sont dangereuses, coûteuses, de longue durée ou très complexes, voire inaccessibles à l'échelle humaine.

Aujourd'hui, la simulation numérique est devenue une démarche essentielle de la recherche scientifique en complément de la théorie et de l'expérimentation. Au niveau national, Genci (Grand équipement national de calcul intensif) est la structure publique chargée de porter la politique française en matière de calcul intensif pour la recherche académique. Avec le ministère de l'Enseignement supérieur et de la Recherche, Genci associe les principaux acteurs français du calcul intensif : le CEA, le CNRS, les universités et l'Inria. «*En quatre ans, les investissements réalisés par Genci ont permis de multiplier par un facteur supérieur à 30 la puissance de calcul mise à la disposition de la communauté scientifique française, qui est actuellement de l'ordre de 600 teraflops^(*)*», poursuit Catherine Rivière.

Au-delà de la France, l'Europe du calcul intensif se concrétise. Convaincus qu'aucun pays ne pouvait, à lui seul, financer et

faire évoluer de manière durable une infrastructure de calcul de visibilité mondiale, vingt représentants de pays européens, dont Genci, ont créé, le 9 juin 2010 à Barcelone, en Espagne, l'infrastructure européenne de recherche Prace (Partnership for Advanced Computing in Europe). Objectif? Mettre en place et animer sur le Vieux continent une infrastructure distribuée et pérenne de calcul, composée de quatre à six centres équipés de machines dotées d'une puissance supérieure au pétaflops^(*).

Laboratoire virtuel

«*Le succès de Prace dépend des résultats scientifiques qui seront obtenus et qui devront être reconnus parmi les meilleurs au monde*», souligne le physicien britannique Richard Kenway, président du conseil scientifique de Prace. *C'est là un objectif fondamental pour nous. La démonstration de notre réussite est un*



DR



PAR
LÆTITIA
BAUDIN
responsable de la
communication
de Genci (Grand
équipement
national de
calcul intensif).



LE FLOPS
(Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un téraflopps permet de faire mille milliards d'opérations par seconde (10^{12}), un pétaflopps d'en faire un million de milliards par seconde (10^{15}) et un exafopps permet d'atteindre le milliard de milliards d'opérations par seconde (10^{18}).

Catherine Rivière,
PDG de Genci.

pré-requis pour élargir le nombre d'États membres de Prace qui contribueront de manière substantielle au fonctionnement de l'infrastructure de recherche. Ces nouvelles contributions permettront à Prace de mettre à disposition des ressources toujours plus compétitives qui favoriseront à leur tour la production d'une meilleure science... C'est un cercle vertueux. » Prace est sur la bonne voie : la Hongrie a rejoint l'infrastructure de recherche, le 8 juin dernier, devenant ainsi le 21^e État européen membre de Prace.

Dès la mi-2010, les scientifiques européens ont eu accès au supercalculateur Jugene, première composante de l'infrastructure Prace, situé à Juelich, en Allemagne. Depuis le début de cette année, ils peuvent également calculer sur Curie, installé en France au Très grand centre de calcul du CEA (TGCC). Financé par Genci, ce supercalculateur sera pleinement opérationnel à la fin 2011 et délivrera une puissance d'au moins 1,8 pétaflops. À partir de 2012, ils auront accès à d'autres machines, en Allemagne, en Italie et en Espagne.

Pour Jérémie Bec, premier scientifique français lauréat d'heures « européennes » de calcul, des progrès scientifiques majeurs sont en vue. «*Les supercalculateurs pétafloppiques ouvrent une nouvelle ère de la recherche, celle de l'expérimentation dans un laboratoire virtuel*», se réjouit ce spécialiste des écoulements turbulents basé à l'Observatoire de la Côte d'Azur, à Nice, qui travaille notamment à élucider le rôle que jouent les fluctuations turbulentes dans le déclenchement des précipitations des nuages chauds.

PRACE



CEA



Inauguration officielle de l'infrastructure européenne de recherche Prace (Partnership for Advanced Computing in Europe), à Barcelone (Espagne), le 9 juin 2010.

Le supercalculateur pétaflopique Curie en cours d'installation au Très grand centre de calcul (TGCC) du CEA, à Bruyères-le-Châtel (Essonne).

des problèmes cruciaux du point de vue économique ou sociétal. Dans le domaine de la climatologie, par exemple, la sauvegarde de la Planète passe par une connaissance fine du climat : « Il est indispensable de disposer d'une très grande puissance de calcul pour simuler avec le plus de réalisme possible le passé de notre climat, les conditions actuelles et son évolution future en fonction de différents scénarios, explique Jean Jouzel, vice-président du Groupe d'experts intergouvernemental sur l'évolution du climat (Giec). Avec Curie, nous pourrions envisager des simulations climatiques à une résolution de l'ordre de la dizaine de kilomètres sur l'ensemble de la Planète et sur plusieurs centaines d'années. Cela nous permettra également d'accroître la participation européenne lors des prochains exercices internationaux de simulation du climat. » Ce n'est là qu'un exemple, d'autres sont présentés dans ce numéro.

Beaucoup reste cependant à faire. « Aujourd'hui, nous ne disposons pas réellement d'un niveau intermédiaire entre les supercalculateurs, nationaux et européens, et les machines de laboratoire, souligne Olivier Pironneau, président du Comité

De façon plus générale, selon Alain Lichnewsky, responsable scientifique de Genci : « L'accroissement des capacités des supercalculateurs, installés en France sous l'égide de Genci ou dans le cadre de Prace, permet des résultats novateurs : généralisation des modèles ab initio reposant sur les principes fondamentaux dans les domaines de la chimie et des matériaux ; obtention de données essentielles pour fonder de nouvelles méthodes expéri-

mentales. Avec les progrès de la modélisation, et l'adaptation des codes de calcul aux nouveaux calculateurs, la frontière des connaissances s'établit maintenant en confrontant l'état de l'art en simulation et la nature des problèmes étudiés. »

Vers un système pyramidal

Ainsi les scientifiques peuvent-ils aborder des phénomènes de plus en plus complexes pour apporter des réponses concrètes à

stratégique du calcul intensif (CSCI), chargé de veiller à la cohérence des actions menées en France dans le domaine du calcul intensif. *Développer les moyens des centres universitaires est une priorité; c'est d'ailleurs l'objectif du projet Equip@meso, ou Équipement d'excellence de calcul intensif de Mésocentres coordonnés, qui est coordonné par Genci et qui associe dix partenaires de différentes régions.*

Retenu dans le cadre de l'appel à projets «Équipements d'excellence» mené sous l'égide du Commissariat général à l'investissement, Equip@meso bénéficie d'un financement de 10,5 millions d'euros pour renforcer les moyens de calcul à l'échelle régionale en complément des moyens nationaux. «*Nous allons ainsi accélérer la construction d'une véritable pyramide du calcul intensif autour de ses trois strates géographiques : les moyens de calcul accessibles au niveau européen, les ressources des centres de calcul nationaux et les moyens coordonnés en région*», complète Catherine Rivière.

En attendant l'Exascale

Ce rapprochement avec les universités doit également permettre de mettre en œuvre une offre concertée et étoffée de formation de spécialistes maîtrisant le calcul intensif et la simulation numérique, à l'instar par exemple du master «Modélisation et simulation», mis en place notamment par le CEA, Centrale Paris, l'École polytechnique et l'université de Versailles Saint-Quentin-en-Yvelines (UVSQ).

«*Nous avons effectivement besoin d'un nombre croissant de jeunes scientifiques formés aux technologies de calcul les plus pointues, capables de comprendre, de développer et de maintenir les logiciels nécessaires*», estime Richard Kenway. C'est que le prochain défi à relever, le passage à l'Exascale^(*), vers 2018, sera de taille : «*Ce n'est pas une vue de l'esprit, confirme Olivier Pironneau, pour réussir cette nouvelle transition, de grands chantiers doivent être menés : amélioration des communications dans les puces; développement de logiciels compte tenu du nombre de processeurs qui seront nécessaires pour atteindre le milliard de milliards d'opérations à la seconde; résistance aux pannes, enfin, qui devrait bénéficier de la virtualisation.*»

En France, cet enjeu décisif

est notamment abordé au travers du Laboratoire européen de recherche sur l'Exascale (ECR Lab), créé conjointement par le CEA, Genci, Intel et l'UVSQ. Accueillant une vingtaine de chercheurs, l'ECR Lab prépare et développe les architectures matérielles et logicielles (codes scientifiques et outils de programmation) qui permettront de soutenir le niveau de performance exaflopique. «*La contribution de Genci s'inscrit notamment dans la perspective de préparer la communauté scientifique française à l'arrivée de l'Exascale*», explique Stéphane Requena, responsable technique de Genci.

Mais le champ du calcul intensif et de la simulation numérique ne se résume pas à la recherche académique : «*Ce sont également des outils stratégiques d'un point de vue économique, rappelle Catherine Rivière. Ils sont un élément essentiel de la productivité industrielle, d'une part en permettant de réduire considérablement le temps de conception et de mise sur le marché d'un produit ou d'un service, d'autre part en contribuant fortement à l'innovation et à l'optimisation des étapes de production et de maintenance.*»

Simulation numérique et PME

Si les grands groupes industriels ou financiers – comme Total, EDF, Airbus ou BNP Paribas – ont communément intégré la simulation numérique et le calcul intensif dans leurs schémas de développement, la démonstration reste à faire auprès des PME qui en maîtrisent souvent moins bien les enjeux technologiques, financiers et humains.

D'où l'initiative «HPC-PME», portée par Genci, l'Inria et Oseo. Bâtie en cohérence avec les recommandations du plan France Numérique 2012, cette initiative a été lancée il y a un peu plus d'un an. «*Notre objectif est d'aider les PME à évaluer la pertinence de l'utilisation de la simulation numérique au regard de leur modèle d'activité, en mobilisant les acteurs du calcul intensif les mieux à même de les accompagner dans cette évaluation*», précise Catherine Rivière. À l'été 2011, pas moins de quinze PME ont exprimé leur intérêt pour un accompagnement. Issues de différents secteurs (automobile, aéronautique, média numérique, industrie navale, microélectronique, traitement du signal, etc.), elles sont répar-



De haut en bas : le physicien britannique Richard Kenway président du conseil scientifique de Prace; Olivier Pironneau, président du Conseil stratégique du calcul intensif.

UNIVERSITY OF EDINBURGH

DR

ties sur l'ensemble du territoire national.

À l'heure où l'Asie domine la course aux supercalculateurs, il est crucial de soutenir le développement scientifique et économique national. La simulation numérique en est un des outils : plus que jamais, il est nécessaire de démocratiser le calcul intensif! ●

INSTALLÉ EN JUILLET 2010, LE SUPERCALCULATEUR TERA 100 EST CLASSÉ PARMI LES PLUS PERFORMANTS DE LA PLANÈTE. CE FUT LE PREMIER EN EUROPE À ACCÉDER AU PÉTAFLUP. SON PLUS GRAND ATOUT : LE RENDEMENT ÉNERGÉTIQUE, QUI EST MULTIPLIÉ PAR UN FACTEUR 7 COMPARÉ À SON AÎNÉ LE TERA 10 POURTANT 20 FOIS MOINS PUISSANT !

LE TERA 100 BRILLE CÔTÉ RENDEMENT



CEA CADAM - P. Stroppa/CEA

Aboutissement de deux années de recherche et développement entre des ingénieurs de la direction des applications militaires du Commissariat à l'énergie atomique et aux énergies alternatives (CEA) et du constructeur informatique français Bull, le supercalculateur Tera 100 est le premier ordinateur conçu et fabriqué en Europe à avoir passé la barre symbolique du pétaflops (*). Sa performance mesurée

de 1,05 pétaflops le plaçait en novembre 2010 au 6^e rang des cinquante ordinateurs les plus puissants du monde; il est repassé à la 9^e place au palmarès établi en juin 2011. Avec un rendement (*) remarquable de 83,7 % au test de référence de ce classement, c'est certainement l'un des supercalculateurs les plus généralistes parmi les dix premiers mondiaux.

Fruit de quinze années de travaux en simulation numérique et architecture des grands systèmes



PAR
PIERRE LECA

chef du
département
sciences de
la simulation et
de l'information
à la Direction
des applications
militaires du CEA

ET SOPHIE
HOUSIAUX

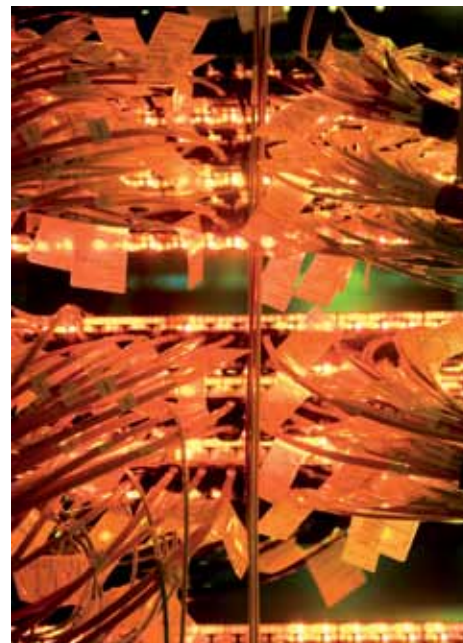
responsable du
projet Tera 100
chez Bull.

de calcul et de gestion des données, le Tera 100 a aussi bénéficié de l'expérience acquise depuis le début des années 2000 avec l'installation et l'exploitation réussies de ses prédécesseurs: le Tera 1 en 2001, puis le Tera 10 en 2005. Grâce à ce savoir-faire et à ces connaissances, tant en termes d'analyse des besoins, d'architecture des calculateurs, que pour l'évolution des technologies matérielles et logicielles, nous avons pu définir les caractéristiques de ce colosse du calcul.

Le Tera 100 a été conçu pour être efficace quelles que soient les méthodes numériques et algorithmiques utilisées. Bâti avec des logiciels libres et des processeurs de grande diffusion, il a donné naissance à une gamme commerciale compétitive de supercalculateurs. En effet, il ne s'agissait pas de réaliser un objet unique, mais d'intégrer ce projet dans une démarche industrielle.

À peine le supercalculateur Tera 100 fut-il installé que l'équipe Bull-CEA envisageait déjà son successeur... 1 000 fois plus puissant !

Le réseau d'interconnexion de la machine possède un câblage complexe qui permet la communication entre les 4 370 nœuds de calcul.



P. Stroppa/CEA

C'est l'architecture très répandue des microprocesseurs X86 qui a été retenue. Le supercalculateur s'intègre ainsi parfaitement à l'environnement de stations de travail fonctionnant sous le système d'exploitation Linux.

Les choix ont tenu compte de l'évolution technologique des microprocesseurs : leur fréquence – ou vitesse de travail – n'augmente guère plus, ce sont leurs cœurs, les unités de calcul élémentaires, qui se multiplient dans chaque microprocesseur. Ceux retenus pour bâtir le Tera 100, Intel Xeon X7560 – connu sous le nom de Nehalem-EX –, comportent huit cœurs. Quatre de ces microprocesseurs sont assemblés pour former un multiprocesseur (le serveur bullx 3060), véritable brique de base du supercalculateur, appelé nœud de calcul dans le jargon des informaticiens. Ainsi, au sein de chaque nœud, 32 cœurs partagent une mémoire de 64 gigaoctets. Pour atteindre sa puissance de calcul, le Tera 100 interconnecte 4370 nœuds, soit 17 480 microprocesseurs et près de 140 000 cœurs.

Consommation maîtrisée

Le réseau d'interconnexion a été conçu selon une architecture originale : une topologie en archipels reliant des grappes de nœuds et permettant un accès rapide aux médias de stockage de données. Le débit maximal agrégé de ce réseau peut atteindre 13 teraoctets (13 milliers de milliards d'octets) par seconde. Globalement, l'architecture du Tera 100 ressemble à un emboîtement de poupées russes : au départ un microprocesseur, puis un nœud de calcul (4 microprocesseurs), puis une armoire (24 nœuds), ensuite un archipel (une vingtaine d'armoiries) et, enfin, le supercalculateur (10 archipels). En outre, le réseau d'interconnexion est à deux niveaux : intra-archipel et interarchipel. Il assure la communication des données entre tout couple de nœuds au sein de l'ordinateur.

Pouvoir maîtriser la consommation d'énergie, problème majeur de ces installations d'envergure, a été l'une des préoccupations essentielles de l'équipe Bull-CEA. Parmi les principales innovations, un dispositif de refroidissement par eau, installé dans la porte des armoires, a permis de rendre l'installation compacte. Le calculateur n'occupe que 650 mètres carrés au sol. Ce

refroidissement, au plus proche du dégagement de chaleur, améliore l'efficacité énergétique du centre de calcul. Par ailleurs, la consommation est modulée en fonction de la charge de calcul, ce qui permet de baisser la fréquence des cœurs s'ils ne sont pas utilisés à pleine puissance. En régime normal, cela devrait limiter la puissance électrique nécessaire à 3 mégawatts. Avec une puissance de calcul 20 fois supérieure à celle du Tera 10, son aîné, le Tera 100 améliore le rendement énergétique d'un facteur 7.

Pour bénéficier des logiciels les plus avancés, les développements ont été menés avec la communauté internationale : certains avec les spécialistes du Department of Energy, aux États-Unis, d'autres avec des industriels et des chercheurs académiques, notamment sur Lustre, le système de gestion de fichiers libre. Ce système dit « distribué » permet de partager les données



LE FLOPS
(Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un téraflopp permet de faire mille milliards d'opérations par seconde (10^{12}), un pétaflopp d'en faire un million de milliards par seconde (10^{15}) et un exafllops permet d'atteindre le milliard de milliards d'opérations par seconde (10^{18}).

Le rendement est le rapport entre la puissance mesurée d'un ordinateur et sa puissance théorique.

réparties sur des centaines de nœuds de calcul. L'environnement logiciel prend en compte la spécificité des nœuds de calcul et la topologie particulière du réseau d'interconnexion. Évidemment, c'est le cas pour la bibliothèque de communication, mais aussi pour le logiciel de gestion de ressources qui s'efforce notamment de placer les calculs selon leur profil.

Enfin, un logiciel d'administration et de supervision contrôle l'état des différents éléments (mémoires, processeurs, réseau...) pour prévenir la conséquence de pannes sur les calculs en cours. Ce maintien en condition opérationnelle est critique durant toute la vie de l'ordinateur.

L'étape du pétaflopp étant franchie, l'équipe Bull-CEA se consacre désormais au nouveau Graal du supercalcul : l'exafllops^(*), soit une puissance de calcul 1 000 fois supérieure à celle du Tera 100. ●

UN FRANÇAIS EN 9^E POSITION MONDIALE

Comme chaque année, en juin et en novembre, un palmarès des cinq cents supercalculateurs les plus puissants du monde est établi sur la base de l'exécution d'un calcul étalon, baptisé Linpack. Dans le dernier classement, de juin 2011, tous les ordinateurs – contre seulement 7 en novembre 2010 – dépassaient la barre du pétaflopp. Un cap franchi pour la première fois en juin 2008 par IBM et son Roadrunner, aujourd'hui 10^e au classement. Le Tera 100 est le supercalculateur le plus puissant d'Europe, tout comme l'était son prédécesseur, le Tera 10, en juin 2006. Avec sa 9^e place au palmarès mondial, le Tera 100 se classe brillamment deuxième en termes de rendement (83,7%), autrement dit pour sa fiabilité par rapport à la puissance théorique annoncée. La Chine continue sa progression fulgurante affichant deux machines dans le dernier Top 10, avec néanmoins de faibles rendements. Les États-Unis restent leaders incontestés côté supercalculateurs installés, mais c'est désormais le Japon qui prend – et de loin – la tête du classement avec le K Computer de Fujitsu et ses 8,16 pétaflopps.

LE TOP 10 DES SUPERCALCULATEURS

RANG	SITE – PAYS	NOM	CONSTRUCTEUR	PUISSANCE MESURÉE (en pétaflopps)	RENDEMENT (en %)
1	RIKEN Advanced Institute for Computational Science – Japon	K Computer	Fujitsu	8,16	93,0
2	National Supercomputing Center, Tianjin – Chine	Tianhe-1A	NUDT MPP	2,57	54,6
3	DOE/Oak Ridge National Laboratory, Tennessee – États-Unis	Jaguar	Cray XT	1,75	75,5
4	National Supercomputing Center, Shenzhen – Chine	Nebulae	Dawning	1,27	42,6
5	Tokyo Institute of Technology – Japon	Tsubame-2.0	HP 3000SL	1,19	52,1
6	DOE/Los Alamos National Laboratory, Nouveau-Mexique – États-Unis	Cielo	Cray XE	1,11	81,2
7	NASA/Ames Research Center/NAS – États-Unis	Pleiades	SGI	1,09	82,7
8	DOE/National Energy Research Scientific Computing Center, Californie – États-Unis	Hopper	Cray XE	1,05	81,8
9	Commissariat à l'énergie atomique et aux énergies alternatives (CEA) – France	Tera 100	Bull Bullx	1,05	83,7
10	DOE/Los Alamos National Laboratory, Nouveau-Mexique – États-Unis	Roadrunner	IBM	1,04	75,7

COMMENT DÉTERMINER L'EFFICACITÉ D'UNE MOLÉCULE À VISÉE THÉRAPEUTIQUE AVANT DE LANCER UNE ÉTUDE CLINIQUE ? LA SOLUTION POURRAIT ÊTRE APPORTÉE PAR LES SUPERCALCULATEURS. EN FIN DE PHASE DE TESTS, LES SIMULATIONS QU'ILS PERMETTENT DEVRAIENT ORIENTER LA RECHERCHE VERS DE NOUVEAUX MÉDICAMENTS.

MODÉLISER LES MOLÉCULES DU VIVANT POUR MIEUX SOIGNER

Pour un traitement thérapeutique, il existe plusieurs milliers de molécules potentielles. Voici le casse-tête quotidien des laboratoires pharmaceutiques. Impossible de mener des essais cliniques pour chaque candidat. Aussi, comment prédire quelles biomolécules ont le plus de chances d'être efficaces ? Une solution, considérée il y a encore peu comme de la science-fiction : la simulation numérique. Une machine colossale, un supercalculateur, pourrait devenir l'outil indispensable pour comprendre le fonctionnement des traitements au niveau moléculaire. « Grâce à la puissance de calcul aujourd'hui disponible, il devient possible de simuler le comportement des biomacromolécules – protéines, acides nucléiques, polysaccharides, etc. – dans leur milieu naturel et de mieux comprendre leurs interactions et rôles fonctionnels au sein de la cellule », explique Richard Lavery, chercheur au laboratoire Bases moléculaires et structurales des systèmes infectieux du CNRS, à Lyon. « En remplaçant les modèles atomiques par des représentations simplifiées, on peut même envisager de bâtir un modèle de virus – contenant l'équivalent de 15 millions d'atomes – et de voir évoluer sa structure dans le temps. »

La simulation permet de sonder ce qui se produit au sein d'une cellule vivante. Concrètement, il s'agit, d'après des observations empiriques, de modé-

liser des systèmes comportant plusieurs dizaines de milliers d'atomes, de prédire les forces qui s'établissent entre ces atomes – les interactions – et d'estimer avec précision leur comportement au sein de l'organisme. Or, plus une interaction est forte, plus la molécule à l'étude a de chances d'être un agent thérapeutique ou diagnostique efficace. « Mais le problème est que la simulation oblige à adopter des modèles de macromolécules qui obéissent à la mécanique classique de Newton plutôt qu'à la mécanique quantique de Schrödinger, pourtant plus adaptée au monde moléculaire », souligne Richard Lavery.

En effet, pour faire simple, une liaison chimique entre deux atomes correspond à un échange d'électrons, un phénomène essentiellement quantique. Mais c'est une équation très difficile à résoudre car il y a des milliers voire parfois des millions d'électrons à prendre en compte. Et si les interactions atomiques sont importantes, il ne faut pas oublier que les atomes bougent, que les structures atomiques se replient, etc. Ces mouvements complexes dans le temps représentent la partie dynamique de la molécule. « Les biomacromolécules ont des structures flexibles et leur fonctionnement implique des mouvements sur plusieurs échelles de temps, allant de la femtoseconde à la seconde », précise Richard Lavery. Aussi, de très nombreux calculs sont nécessaires pour parvenir à se représenter leur

comportement avec suffisamment de précision.

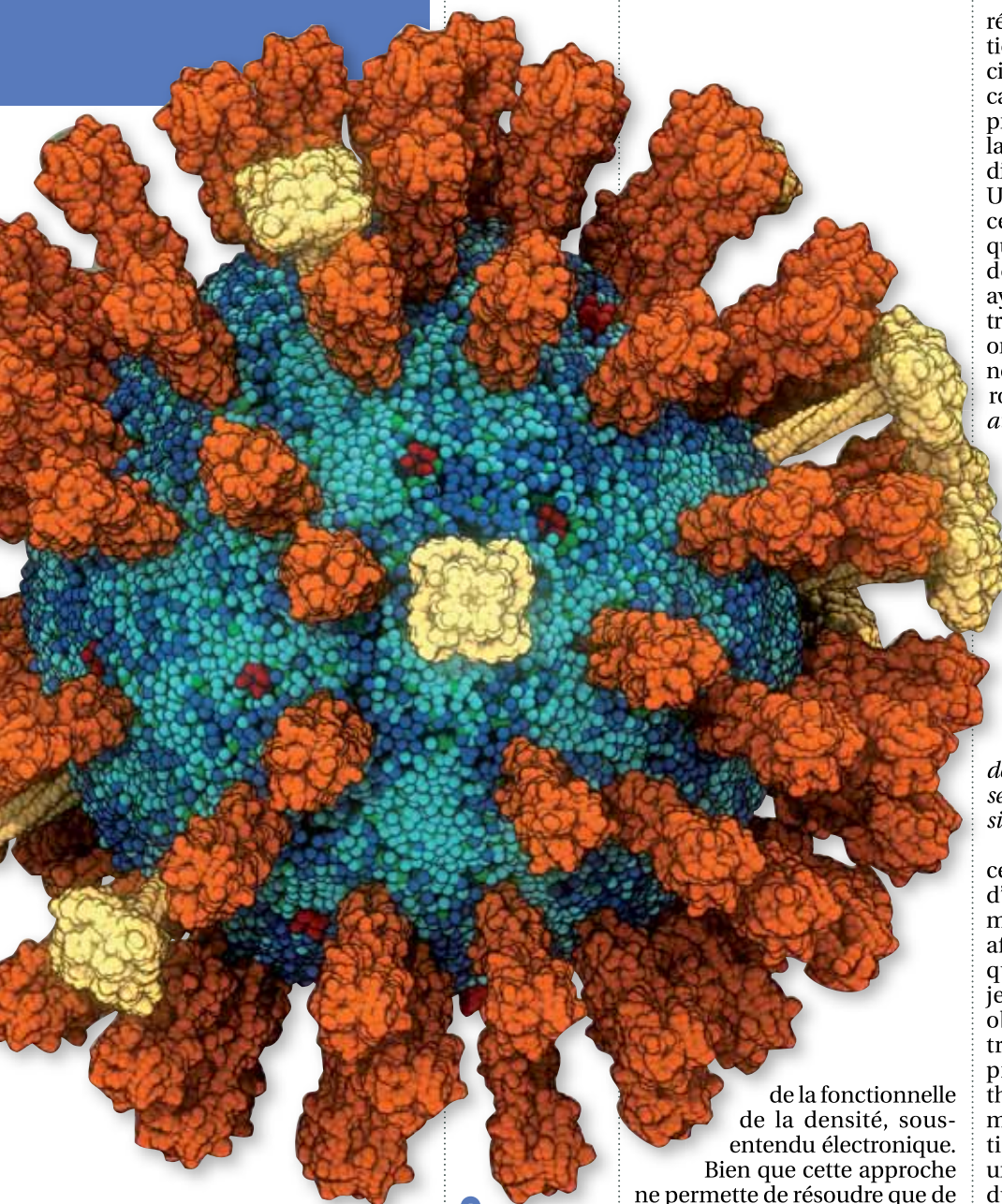
Scalabilité idéale

C'est essentiellement pour cette raison que la puissance des supercalculateurs est capitale, puisqu'elle permet d'effectuer ces calculs en des délais toujours plus courts. « Au début des années 2000, prédire l'intensité d'une seule interaction entre une molécule à visée thérapeutique et une biomolécule demandait trois mois de calculs ! », se souvient Michel Masella, chercheur au laboratoire de Chimie du vivant au Commissariat à l'énergie atomique et aux énergies alternatives (CEA). À l'époque, la simulation numérique faisait sourire de nombreuses personnes qui estimaient que l'on avait plus vite fait de faire des expériences que d'attendre trois mois pour un seul résultat. « Aujourd'hui, notre objectif est de réaliser de 100 à 1 000 prédictions d'interaction par jour, ce qui sera possible grâce aux ressources fournies par le supercalculateur européen Curie, hébergé par le CEA, et qui disposera bientôt 80 000 processeurs. » Un saut phénoménal ! À la fois technique, car il a fallu des machines beaucoup plus performantes, mais également un saut aux niveaux physique et mathématique, car les méthodes que les chercheurs utilisent actuellement se sont incroyablement développées sur le plan algorithmique pour gagner en performances. Certaines unités de recherche du CEA continuent



SOLVATATION

Phénomène physico-chimique observé lors de la dissolution d'un composé chimique dans un solvant, qui voit les atomes, ions ou molécules, de l'espèce chimique se disperser dans la solution en interagissant avec les molécules du solvant.



d'ailleurs de se consacrer à l'amélioration du rendement quantitatif de codes algorithmiques, c'est-à-dire en augmentant le nombre d'opérations réalisées par seconde, ou flops.

En revanche, ces méthodes de calcul quantique « ne passeront pas l'échelle » vers l'exascale. Dans le jargon informatique, on parle de « non-scalabilité », ce qui signifie qu'il est impossible avec ce type de calculs d'exploiter au mieux les potentialités des ordinateurs exascale à venir, comprenant des millions voire des milliards de processeurs. Aujourd'hui, la méthode-phare pour le calcul quantique des molécules du vivant est une méthode ayant émergé dans les années 90 : la DFT (pour *Density Functional Theory*) ou théorie

• Ce modèle de l'enveloppe du virus de la grippe a été obtenu par simulation numérique. Il montre les trois protéines principales (orange, crème et rouge) et permet d'étudier leurs interactions avec les molécules à visée thérapeutique testées. Travaux de D. Parton, M. Sansom (Univ. Oxford) et M. Baaden, M. Chavent, A. Tek (IBPC, Paris).

de la fonctionnelle de la densité, sous-entendu électronique.

Bien que cette approche ne permette de résoudre que de manière approchée l'équation de Schrödinger décrivant la « nature » quantique des électrons, elle est considérée comme un bon compromis entre rapidité et précision chimique. Malheureusement, à cause de sa mauvaise « scalabilité », tirer avantage des plates-formes à plus de quelques milliers de processeurs semble très difficile avec elle.

Trajectoires aléatoires

C'est pourquoi la recherche s'oriente vers le développement de nouvelles approches intrinsèquement adaptées à des machines qui auraient un nombre arbitraire de processeurs. Ainsi, à Toulouse au laboratoire de chimie et physique quantiques, Michel Caffarel dirige une équipe de recherche consacrée au développement d'une méthode de

résolution alternative de l'équation de Schrödinger, aussi précise que la DFT, mais passant le cap d'un nombre quelconque de processeurs – on parle de « scalabilité idéale ». Une méthode dite de Monte-Carlo quantique. Un nom qui fait référence au célèbre casino monégasque et qui s'explique par l'introduction de trajectoires électroniques ayant un caractère aléatoire, ces trajectoires étant construites sur ordinateur en tirant des séries de nombres au hasard comme à la roulette du casino. « L'avantage, avec la dynamique moléculaire probabiliste des électrons, c'est qu'elle est parallélisable : on peut effectuer les calculs en parallèle sans avoir à connaître ce qu'il se passe sur chacun des processeurs », précise Michel Caffarel. Pourquoi ? Tout simplement parce que la simulation peut être découpée à loisir en un ensemble de trajectoires électroniques indépendantes. Une propriété unique pour exploiter un nombre arbitraire de processeurs, puisque ceux-ci ne se parlent pas tout au long de la simulation. »

En pratique, chaque processeur se voit confier le calcul d'une trajectoire. Reste à faire la moyenne des résultats obtenus afin de reconstituer le résultat qui correspondrait à une trajectoire électronique unique, obtenue par juxtaposition des trajectoires individuelles. Au printemps dernier, cette méthode a fait ses preuves sur la machine Curie. En collaboration avec Anthony Scemama, un jeune ingénieur de recherche du CNRS, une simulation a pu être effectuée sur l'ensemble des 10 000 processeurs de la machine avec une scalabilité parfaite. La simulation a ainsi pu être réalisée en un temps 10 000 fois inférieur à celui qu'aurait demandé le même calcul sur un seul processeur.

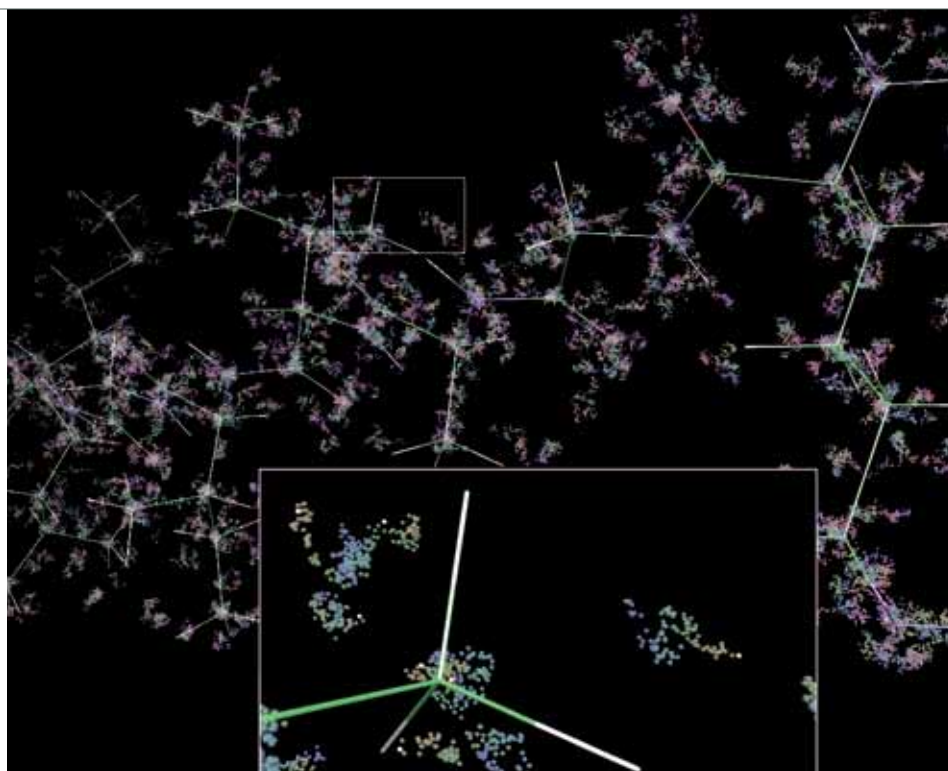
« Cela fait plus de vingt ans que je développe ces méthodes et je pense qu'on arrive à un tournant : on pourrait supplanter d'ici quelques années les méthodes actuelles ! », s'enthousiasme Michel Caffarel. Et la phase de validation de commencer. « En octobre, nous ferons une étude sur l'interaction de molécules à la base de la chimie de la maladie d'Alzheimer – l'agrégation des peptides amyloïdes – en utilisant les 80 000 processeurs de Curie, et il semble envisageable à terme de contribuer à la compréhension de

certains aspects des maladies neurodégénératives.» Un vrai défi! Et surtout l'espoir de développer la recherche sur des pathologies pour lesquelles les essais cliniques sont très durs à réaliser.

Plus loin, c'est la perspective d'une recherche médicale ciblée qui se dessine. «*Nous allons commencer à intégrer les simulations numériques précises en amont de tous les projets expérimentaux, ce qu'on ne faisait avant qu'avec des modèles très, voire trop simplifiés, pour entrer dans une phase prédictive d'une grande efficacité*», prévoit Michel Masella. C'est la coopération entre théorie et expérimentation qui est en train de se mettre en place: elle devrait permettre aux chercheurs de prédire des voies à forte probabilité de succès.

Histoire d'interactions

«*Mais il faut rester réaliste, tempère cependant Michel Masella. En biologie, les réactions chimiques se jouent parfois à presque rien.*» Bien que cette approche soit très prometteuse, il faut donc continuer à optimiser les molécules «*présélectionnées*». D'où la nécessité d'affiner au maximum les modèles. «*On corrige petit à petit tous les défauts connus. Par exemple, nous travaillons en ce moment sur les interactions électrostatiques à longue portée. Car les modèles actuels n'analysent des interactions qu'à des distances assez petites, ce qui décrit assez mal la solvation des molécules*



chargées. Or, c'est un paramètre important pour comprendre la réaction du médicament dissout dans un solvant. Désormais, le but est d'adapter le modèle aux interactions chargées.»

En outre, si l'on veut comprendre au mieux ce qui se produit dans une cellule vivante, il ne faut pas oublier qu'aux interactions thérapeutiques du médicament avec une protéine cellulaire s'ajoute le fait que les protéines interagissent également les unes avec les autres. Et, pour rendre les choses encore

➊ À chaque pas de simulation, les électrons apparaissent d'une couleur différente (rouge -> violet -> bleu -> vert -> jaune) pour permettre de voir l'évolution de la position des électrons au cours du temps.

plus complexe, la génétique fait que chaque individu peut présenter des variations au niveau de ces protéines. Si parvenir à une simulation individualisée d'interactions thérapeutiques relève donc pour l'heure de la science-fiction, Michel Masella n'en rejette pas l'hypothèse: «*Lorsque nous serons capables de bien identifier ces mutations, nous pourrons alors entrer dans ce que l'on appelle de la personnalisation. Mais pour cela, il faudra bien attendre encore vingt ans!*» ➋

MORGANE KERGOAT

UNE EXIGENCE DE PRÉCISION

Contrairement à l'aéronautique ou à la climatologie, qui utilisent les modèles numériques depuis longtemps, la biochimie ne les intègre que difficilement dans sa culture. En effet, dans les deux premiers domaines, il n'est pas besoin d'avoir des modèles très évolués pour commencer les prédictions par simulation. Tandis qu'en biologie il est exigé qu'en phase de test, un résultat corrobore totalement l'expérience pour que la méthode puisse être validée. Si ce n'est pas le cas et qu'il y a un décalage, les résultats ne sont pas à corriger; c'est l'ensemble du modèle théorique qui doit être repris. Ceci exige un degré de précision de calcul extrême et une adaptation constante des codes. Une culture de l'excellence, qui a pour revers de creuser l'écart avec les autres domaines scientifiques ayant plus largement recours aux simulations numériques.



LORS D'UN SÉISME SOUS-MARIN PROVOQUANT UN TSUNAMI, LES TEMPS DE CALCUL DES HAUTEURS D'EAU ATTENDUES SUR LES CÔTES EMPÊCHENT TOUTE UTILISATION EN TEMPS RÉEL. AVEC LES SUPERCALCULATEURS, LES EFFETS DES TSUNAMIS POURRAIENT ÊTRE PRÉVISIBLES SUR LES CÔTES EN QUINZE MINUTES.

LE SUPERCALCULATEUR AU SERVICE DE L'ALERTE TSUNAMI

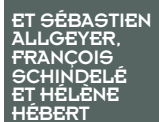
Nous l'avons encore vu en mars dernier au Japon, les tsunamis sont capables de dévaster des côtes entières en provoquant des destructions considérables. Ces vagues particulières se traduisent par des inondations successives de la côte – toutes les 20 à 40 minutes –, alternant avec des retraits marqués de la mer. Les tsunamis sont générés par de très forts séismes, d'une magnitude généralement supérieure à 7,5 sur l'échelle de Richter, qui se produisent sur les zones de subduction (*). C'est dans le Pacifique que ces phénomènes sont les plus fréquents, du fait de l'activité tectonique intense dans cette partie du Globe.

Durant les années 1960, à la suite de cinq tsunamis catastrophiques sur les côtes du Pacifique, le Centre polynésien de prévention des tsunamis (CPPT) a été créé au Laboratoire de géophysique du CEA basé à Tahiti. Sa mission est d'assurer une veille permanente de l'activité sismique globale de l'océan Pacifique afin de signaler toute alerte aux autorités du territoire polynésien. La région la plus proche où peuvent se produire des séismes se situe dans la zone des Tonga – Kermadec; le temps de parcours minimal d'un tsunami généré dans cette région pour at-



PAR ANTHONY JAMELOT ET DOMINIQUE REYMOND

du Laboratoire de Géophysique de Pamatai à Tahiti



ET SÉBASTIEN ALLGEYER, FRANÇOIS SCHINDELE ET HÉLÈNE HÉBERT
du département Analyse, surveillance, environnement de la Direction des applications militaires du CEA.

teindre les côtes polynésiennes est de trois heures.

Lorsqu'un séisme est détecté, les chercheurs analysent sa localisation et sa magnitude puis estiment s'il est susceptible de conduire à la formation d'un tsunami. L'estimation de l'amplitude du tsunami attendu était traditionnellement réalisée selon une loi empirique obtenue à partir des précédents événements observés en Polynésie. C'est ainsi qu'a été élaborée l'alerte au tsunami qui est survenu au Chili en février 2010. Une autre méthode, dont le développement a débuté en 2009 au CPPT, a pu être utilisée de façon opérationnelle lors du tsunami du Japon en mars 2011. Elle s'appuie sur une base de données de 260 scénarios de tsunamis pré-calculés, dont les sources fictives sont réparties le long des principales zones de subduction du Pacifique. Cette nouvelle méthode permet de diminuer l'incertitude sur les hauteurs d'eau attendues, et ce, très rapidement après le séisme, mais ne donne cependant pas les cartographies détaillées des hauteurs attendues à la côte.

Une triple modélisation

Plus précise, la simulation complète des tsunamis jusqu'à un ni-

veau détaillé sur la côte demandait jusqu'à récemment un temps de calcul prohibitif pour une utilisation en temps réel. La parallélisation du code sur les calculateurs haute performance ouvre cependant de nouvelles perspectives.

Les simulations numériques des tsunamis reposent sur la modélisation de trois phénomènes : les déformations du plancher océanique causées par le séisme, la propagation en océan profond et les effets côtiers. La déformation initiale de la surface de l'océan est calculée par des modèles élastiques de déformation de la croûte terrestre. La propagation au large, quant à elle, repose sur la résolution des équations non linéaires de la mécanique des fluides dans l'hypothèse des « ondes longues » : les longueurs d'onde des tsunamis (100 à 300 km) sont largement supérieures à la profondeur du milieu de propagation (4 à 6 km). Enfin, la simulation des effets du tsunami à l'approche des côtes n'est possible que si des données bathymétriques (profondeurs marines) et topographiques de haute résolution sont disponibles. L'acuité de la représentation des processus physiques tels que l'inondation, les tourbillons



La hauteur d'un tsunami lorsqu'il arrive sur une côte (ci-contre, photo montage) n'est aujourd'hui mesurée qu'après son passage. Des calculs plus rapides permettront de l'anticiper et d'évacuer les populations.

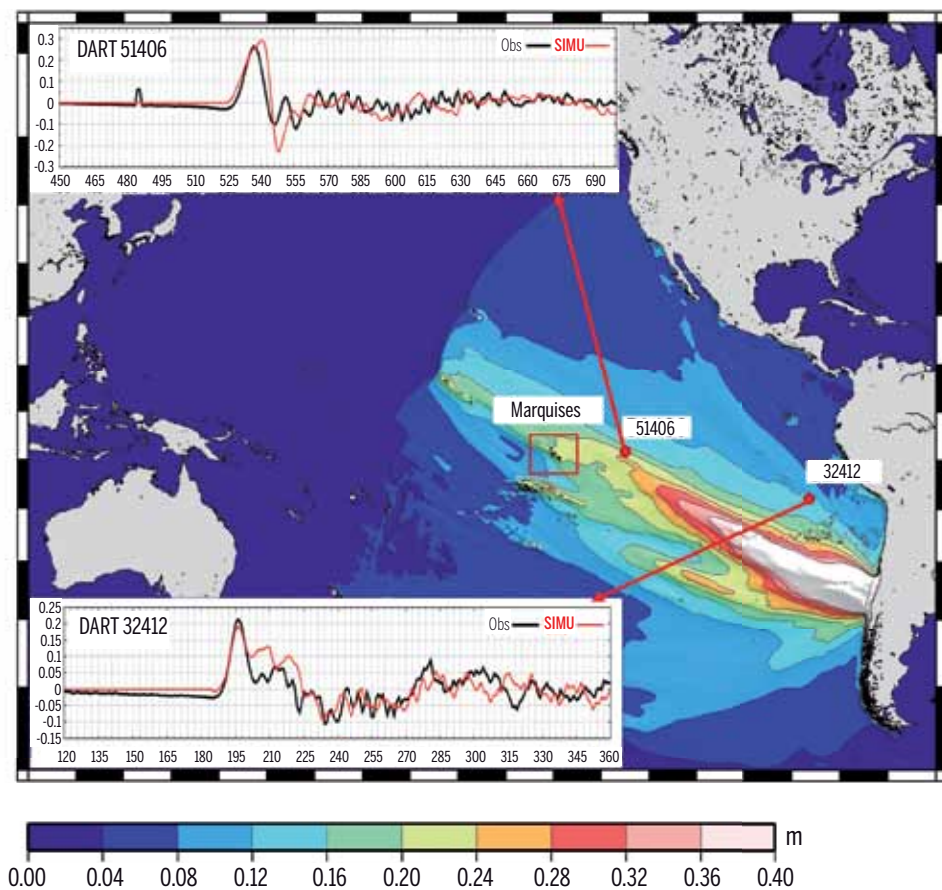


Figure 1. Hauteurs maximales après quinze heures de propagation du tsunami. Comparaison des enregistrements de deux tsunamimètres DART® (Deep-ocean Assessment and Reporting of Tsunamis) et simulés.

moins de 15 minutes de calcul, alors qu'il aurait fallu environ 36 heures sur un seul processeur. Ils ont permis d'établir la distribution des hauteurs d'eau maximales au large après quinze heures de propagation – soit environ 10 000 km parcourus. On a ainsi pu voir que la Polynésie était dans l'axe d'énergie principale (Figure 1).

Estimer les hauteurs d'eau

D'autre part, la simulation numérique du tsunami propagé dans des grilles de plus en plus fines au voisinage des côtes (10 à 15 m de résolution) a permis d'estimer la distribution des hauteurs d'eau maximales, ainsi que les champs de vitesses horizontales décrivant les courants à un instant donné pour les quinze baies. Conclusion : les marégrammes synthétiques sont bien comparables aux observations réelles des marégraphes portuaires ; de même, la comparaison est cohérente avec les hauteurs d'eau rapportées dans des témoignages ou sur des photographies, telle la baie de Tahauku sur l'île d'Hiva Oa. Un tourbillon photographié dans la baie de Hakahau (île de Ua Pou), après environ 11 heures 45 minutes de propagation, est reproduit au même instant.

Ce type de simulation complète du tsunami, de sa source jusqu'aux côtes, montre à quel point la précision des résultats constitue un outil primordial dans le cadre de l'alerte, et complémentaire aux méthodes existantes. En effet, connaître à l'avance le niveau d'inondation sur les côtes avec une incertitude minimale révolutionnerait la gestion de ladite alerte. Ce n'est pas encore possible aujourd'hui, mais on peut envisager dans un avenir proche d'obtenir, en moins d'une heure, les résultats de ces simulations prédictives pour toute la Polynésie française. En supposant bien sûr que des moyens de calculs adéquats soient dédiés au système d'alerte. ●

ou encore l'amplification par résonance dans un port dépend de la résolution des grilles de calcul. Avec le modèle que nous utilisons, nous accédons à une description de la bathymétrie avec une résolution spatiale allant de 5 km pour tout le Pacifique et jusqu'à 15 mètres voire 10 mètres dans les ports et les baies. Nous disposons de données bathymétriques et topographiques de haute définition pour dix-neuf baies polynésiennes. La parallélisation du code et l'utilisation du Centre de calcul recherche et technologie (CCRT) du CEA Île-de-France ont été incontournables pour multiplier les études et diminuer les incertitudes.

Mais que valent ces techniques, confrontées à l'expérience ? Nous avons pu l'évaluer à la suite du tsunami du Chili, le 27 février 2010. À 6 h 34 mn GMT s'est produit un violent séisme de magnitude 8,8 près des côtes chiliennes. Comme attendu, vu la force du séisme, le tsunami a été très destructeur au Chili et s'est ensuite propagé à travers tout le Pacifique. En Polynésie française, ce sont les îles Marquises qui ont été les plus touchées. En effet, les pentes sous-marines peu raides et les larges baies ouvertes sur l'océan sans récif protecteur fa-

vorisent l'amplification des tsunamis. Depuis le milieu du XIX^e siècle, plus de quinze tsunamis ont ainsi été observés dans les baies de ces îles.

De 36 heures à 15 minutes

Étant donnée la magnitude du séisme du Chili, la loi empirique utilisée lors de cette alerte indiquait des hauteurs pouvant atteindre 3 m aux Marquises et 2 m à Tahiti. Le niveau d'alerte est alors passé rapidement au rouge, ce qui implique l'évacuation des zones littorales et des ports. Les niveaux d'eau maxi (« crête à creux ») mesurés sur les marégraphes portuaires ont finalement atteint plus de 3 m à Hiva Oa et Nuku Hiva, contre seulement 35 cm environ dans le port de Papeete. Les hauteurs mesurées aux Marquises sur les côtes (hors marégraphes) vont jusqu'à 3 m au-dessus du niveau des plus basses mers, mais le tsunami est arrivé à marée basse. L'alerte a bien fonctionné, il n'y a eu aucune victime ; seuls quelques dégâts à déplorer – une embarcation dont le propriétaire avait refusé l'évacuation. La simulation numérique de cet événement a été réalisée *a posteriori* sur deux cents processeurs du CCRT, pour quinze baies des îles Marquises. Les résultats ont été obtenus en



SUBDUCTION

Phénomène au cours duquel une plaque lithosphérique plonge sous une autre plaque et s'enfonce dans le manteau terrestre.

AUJOURD'HUI, L'INDUSTRIE NUCLÉAIRE NE POURRAIT PLUS S'ENVISAGER SANS LA SIMULATION. ELLE PERMET D'OPTIMISER TOUT LE CYCLE DU COMBUSTIBLE, JUSQU'À LA GESTION DES DÉCHETS. UNE MODÉLISATION TRIDIMENSIONNELLE AU SERVICE DES INGÉNIEURS... ET DE LA SÛRETÉ DES RÉACTEURS.

LES FUTURS RÉACTEURS NUCLÉAIRES PROFITENT DÉJÀ DU CALCUL HPC

Que ce soit pour la sûreté, l'extension de la durée de fonctionnement des réacteurs, ou l'optimisation de la gestion des déchets, la simulation joue un rôle capital et croissant dans l'industrie nucléaire. Les équipes d'ingénierie comme de R&D l'utilisent de plus en plus : pour calculer le comportement normal des systèmes bien sûr, mais aussi pour imaginer des domaines de fonctionnement au-delà de ce que l'expérience peut mesurer. Un point majeur pour la sûreté nucléaire. Comme dans tout secteur, cette approche repose sur un triptyque de modélisation des phénomènes physiques, de simulation numérique et de validation expérimentale.

Dans le domaine de la physique des réacteurs et du cycle du combustible, plusieurs phénomènes physiques font l'objet de calculs. La cinétique et la répartition des neutrons dans le cœur déterminent le contrôle de la réaction en chaîne et la maîtrise du combustible nucléaire. La propagation des rayonnements ionisants est calculée à la fois pour la protection des personnes et pour connaître les effets sur les matériels. Enfin, l'évolution du combustible nucléaire est en lien direct avec l'optimisation de l'utilisation des ressources en matières fissiles et la gestion des déchets.

La modélisation théorique de ces phénomènes repose sur les deux équations de bilan neutroniques : l'équation de Boltzmann permet de modéliser la vie des neutrons, tandis que l'équation de Bateman traduit l'évolution des isotopes au cours du temps. Ce sont deux équations « exactes », c'est-à-dire « sans approximation ». Elles utilisent des grandeurs physiques qui caractérisent l'interac-



PAR CHRISTOPHE CALVIN
chef de laboratoire au sein du Service d'études des réacteurs et de mathématiques appliquées à la Direction de l'énergie nucléaire du CEA, expert en calcul haute performance appliqué à la physique des réacteurs.

tion des particules, comme les neutrons, avec le combustible et les matériaux du réacteur. Et les données nucléaires de base sont issues de mesures expérimentales précises.

Finesse de modélisation

Tout cela semble parfait : des données précises et des équations qui permettent de modéliser au plus près les phénomènes physiques. Néanmoins, ces équations imposent de résoudre des systèmes qui comportent plus de... 1 000 milliards

d'inconnues ! Mission impossible, quelle que soit la puissance de calcul. Aussi, afin de résoudre ces équations, on fait appel à deux approches complémentaires. D'une part l'approche déterministe : elle repose sur des hypothèses physiques et des modèles numériques afin de résoudre le problème (ce sont les codes APOLLO2 et APOLLO3®, développés à la direction de l'énergie nucléaire du CEA, CEA/DEN). D'autre part l'approche Monte-Carlo : elle fait appel à la représentation native des données



Le cœur d'un réacteur nucléaire est chargé avec du combustible. Sa composition et son positionnement sont des paramètres clés pour la puissance et la sûreté d'une centrale.

CEA

pour «jouer» la vie de milliards de neutrons «utiles» (code TRIPOLI-4TM, également développé au CEA/DEN). Ces méthodes de simulation numérique sont étroitement couplées à des expériences et des mesures.

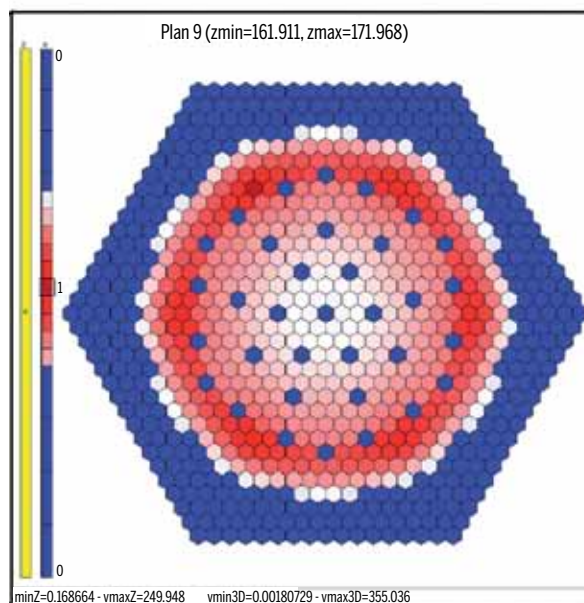
Dans le domaine de la physique des réacteurs, le calcul haute performance (HPC) conduit à modifier la façon d'utiliser la simulation numérique. Les évolutions des codes calcul (algorithmes, méthodes numériques, etc.) et l'utilisation efficace de puissances de calculs croissantes font tout d'abord continuellement progresser la précision et la finesse des modélisations. Les informations ainsi obtenues par la simulation numérique sont aussi plus complètes grâce à la généralisation des calculs tridimensionnels – apparus au début des années 1990 – et la prise en compte des phénomènes multi-physiques (neutronique, thermo-hydraulique, modélisation du combustible, etc.). Autre avancée : la possibilité de simuler de manière simultanée un nombre toujours plus grand de composants des centrales. Par exemple, le HPC ouvre la voie à la modélisation, au sein d'une même simulation et en trois dimensions, du cœur du réacteur et de la chaudière en prenant en compte de manière couplée les phénomènes neutroniques et thermo-hydrauliques – réseaux de circulation primaire de l'eau, générateurs de vapeur, etc. Les ingénieurs obtiennent ainsi des résultats beaucoup plus précis sans consacrer plus de temps qu'ils ne le faisaient avec des modèles simplifiés utilisés couramment.

Prenons l'exemple du code de simulation neutronique APOLLO3, qui a été utilisé à la fin 2010 sur la nouvelle machine de calcul pétasflopique (Tera-100) du CEA/DAM, pour réaliser des simulations tridimensionnelles d'un réacteur de quatrième génération, successeur de l'EPR. La détermination des incertitudes et des biais des schémas de calcul des réacteurs nucléaires est un des points clés pour la conception des prochains réacteurs. En associant le calcul haute performance à un code moderne, les ingénieurs parviennent à une évaluation efficace de ces incertitudes. Ils peuvent

Sur cette représentation d'un cœur de réacteur nucléaire de quatrième génération, calculée à l'aide du code APOLLO3, chaque losange représente un assemblage de combustible nucléaire. Le niveau de puissance neutronique de chaque assemblage est représenté par une couleur, du bleu (puissance nulle) au rouge (puissance maximale). Les losanges bleus au centre sont les barres de contrôle servant à piloter la réaction nucléaire.



Le flops (Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un téraflops permet de faire mille milliards d'opérations par seconde (10^{12}).



connaître beaucoup plus rapidement l'impact précis d'un maillage de calcul plus ou moins fin sur le résultat final. Ils déterminent ainsi des solutions qualifiées de type «référence» pour répondre aux besoins d'évaluation des incertitudes. Et pour cela, on déploie les grands moyens : des puissances de calcul de plusieurs centaines de téraflops^(*), utilisant de manière simultanée plus de 30 000 processeurs. À l'arrivée : un calcul du cœur du réacteur en trois dimensions extrêmement réaliste, grâce à la résolution exacte de l'équation du transport des neutrons obtenu en seulement quelques heures. Avec un unique processeur, il aurait fallu au bas mot une année entière...

Chargement du cœur

Le calcul haute performance permet enfin de mettre en œuvre des méthodes novatrices d'optimisation des paramètres de fonctionnement et de maîtriser de manière systématique les incertitudes. Un exemple caractéristique de ce type d'approche est l'utilisation combinée de la puissance de calcul, des méthodes d'intelligence artificielle et de codes de simulation de nouvelle génération pour optimiser des plans de chargement de cœur de réacteur. De quoi s'agit-il ? Un cœur de réacteur nucléaire est constitué d'assemblages combustibles. Ces assemblages restent un certain temps dans le cœur

du réacteur. Leur positionnement en fonction de leur type (nature du combustible et temps de présence dans le cœur) constitue le plan de chargement du cœur. Selon ce plan de chargement, des paramètres majeurs de fonctionnement en termes de sûreté et de rentabilité – comme la puissance maximale du réacteur – peuvent être optimisés.

Jusqu'à alors, ces optimisations étaient réalisées grâce à l'expertise humaine et au retour d'expérience sur les réacteurs existants. Mais sur les nouveaux concepts de réacteur, il est de plus en plus difficile et surtout très long d'optimiser «à la main» le chargement de combustible, vu le nombre de configurations possibles.

La conception d'un outil d'optimisation adapté à la détermination des plans de chargement des combustibles nucléaires et indépendant de la configuration étudiée – cœur, combustible – permet de faciliter le travail des ingénieurs et surtout de diminuer les temps d'études. Cet outil, basé sur un logiciel d'optimisation multicritères par algorithme génétique (VIZIR) et sur le code de calcul neutronique APOLLO3[®], permet au concepteur d'améliorer les performances et la sûreté des réacteurs. C'est ainsi qu'avec l'aide de 4 000 processeurs on peut désormais trouver des solutions de plan de chargement de cœurs de réacteurs complexes en moins de 24 heures. ●

COMMENT LES ATOMES SE LIENT-ILS ENTRE EUX POUR FORMER DES MATÉRIAUX ? SIMULER LA CROISSANCE AU NIVEAU ATOMIQUE RESTE UN GRAND DÉFI. L'OBJECTIF EST DE MIEUX LA MAÎTRISER EN FONCTION DU FLUX D'ATOMES, DE LA TEMPÉRATURE, DU CHAMP ÉLECTRIQUE... UNE COMPRÉHENSION PRIMORDIALE DANS LE DOMAINE DE LA NANOÉLECTRONIQUE.

VOIR LES MATÉRIAUX GRANDIR ATOME PAR ATOME

Au niveau atomique, la croissance d'un matériau est un mouvement des atomes arrivant à la surface et s'incorporant aux atomes déjà présents. Et ce, en minimisant l'énergie du système, raison pour laquelle les atomes s'organisent en réseaux très réguliers, comme les cristaux. Pour étudier ce phénomène, il faut d'abord identifier les différents mécanismes atomiques de diffusion, puis les mettre en compétition entre eux en fonction de la température. Cela requiert de pouvoir simuler la matière à l'échelle atomique.

On peut modéliser la croissance à partir de modèles phénoménologiques : un modèle avec peu de paramètres et reposant sur certaines hypothèses permet de décrire l'essentiel du phénomène physique. Mais il est difficile de savoir si un tel modèle rendra bien compte des phénomènes et s'il aura des capacités prédictives sans faire des expériences réelles ou numériques.

Aussi, le mieux est de faire appel à des méthodes dites de « premiers principes », ou *ab initio*, qui s'appuient sur l'équation de Schrödinger de la mécanique quantique, capable de prédire de façon fine les configurations atomiques les plus stables. Elles permettent de voir les événements rares qui gouvernent la croissance, comme le mouvement ou la diffusion d'un atome d'un site stable à un autre.

Barrières d'énergie

En pratique, il faut coupler une méthode de calcul de la structure électronique comme le code BigDFT (développé depuis 2005 par le laboratoire L_Sim du CEA, à Grenoble), avec un algorithme de recherche de *minima* et de



PAR THIERRY DEUTSCH

chercheur à l'INAC (Institut nanosciences et cryogénie) au CEA de Grenoble, où il dirige le laboratoire de simulation atomistique



ET PASCAL POCHET

chercheur à l'INAC.

barrière d'énergie le plus exhaustif possible comme MH (« Minima Hopping », inventé par le groupe de Stefan Goedecker à Bâle) ou bien ART (« Activation Relaxation Technics », développé par Normand Mousseau, de l'université de Montréal).

Actuellement, de nombreux algorithmes sont mis au point et testés pour chercher le plus rapidement possible les configurations atomiques les plus stables et les barrières d'énergie. Ces algorithmes sont nécessaires pour comprendre entre autres le repliement des protéines.

La seconde étape consiste à mettre en compétition les différents mécanismes de diffusion des atomes. Pour cela, on fait appel à la physique statistique, et plus particulièrement à un algorithme de Monte Carlo cinétique, qui consiste à « tirer » aléatoirement une nouvelle position atomique en fonction de la barrière d'énergie à franchir. On peut ainsi réaliser des expériences numériques de croissance des matériaux en considérant les atomes un par un.

Mais c'est le calcul des *minima* et des barrières d'énergies en utilisant des méthodes *ab initio* qui est le facteur limitant de cette approche, à cause de l'immense puissance de calcul nécessaire. En effet, il faut ici résoudre l'équation de Schrödinger ! On fait appel au « formalisme de Kohn-Sham », qui repose sur le « théorème de la fonctionnelle de la densité » ayant donné lieu au prix Nobel de chimie en 1998. De quoi s'agit-il ? L'idée sous-jacente est que la densité électronique suffit à déterminer l'état fondamental des électrons pour une position atomique définie. En considérant que les électrons sont toujours à l'équilibre lorsque les atomes bougent, il est alors pos-

sible de calculer des forces atomiques et de trouver les *minima* – c'est-à-dire les configurations atomiques stables – et les barrières d'énergie.

Notre groupe s'est intéressé à la simulation de la croissance du graphène sur carbure de silicium (SiC) (Figure 1). Nous avons considéré des surfaces périodiques dans deux directions constituées de l'ordre de 700 atomes, soit 2608 électrons. Dans l'une de ces directions, le cristal de carbure de silicium (SiC) est constitué d'une alternance de plan pur carbone et silicium. La surface peut donc se terminer par un plan carbone ou un plan silicium. Les méthodes *ab initio* sont ici incontournables pour décrire notamment les liaisons des atomes de carbone, différentes dans le carbure de silicium et dans le graphène. Pour donner un ordre, le calcul d'une configuration atomique requiert environ cinq heures de calcul en utilisant 600 processeurs mis en parallèle dans un supercalculateur !

Les ondelettes en renfort

Le code BigDFT utilise de manière originale de nouvelles fonctions mathématiques, les ondelettes, jusque-là principalement utilisées pour la compression d'images. Il a été optimisé – en partenariat avec le laboratoire d'informatique de Grenoble – pour utiliser plusieurs cœurs de calcul par électron simulé. En pratique, il est donc possible pour notre système de 2608 électrons d'utiliser plus de 2608 cœurs de calcul. Le code BigDFT est aussi capable d'utiliser des processeurs graphiques, avec un temps de calcul encore divisé d'un facteur 10.

Le supercalculateur Tera 100 permet de réaliser des calculs 2000 fois plus rapides qu'un ordinateur classique. Nous arrivons ainsi à

FIGURE 1

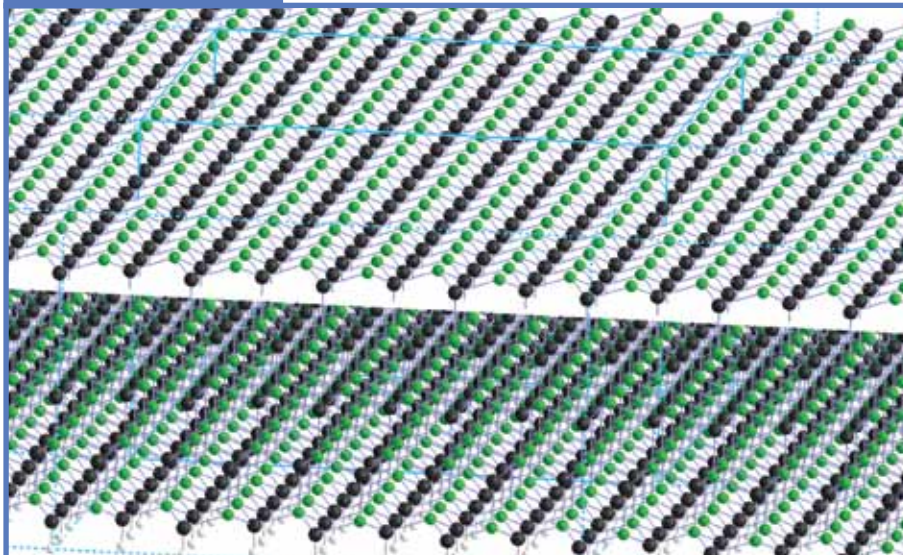


Figure 1

Surface nue de SiC terminée silicium (en vert) au-dessus du plan jaune. La surface du bas est terminée carbone (en noir). La boîte délimitée par les traits bleus correspond à la boîte de calcul contenant 700 atomes.

Figure 2

Vue du dessus de la surface de SiC avec une nanofeuille composée de 16 atomes de carbone de couleur bleue. Les liaisons en rouge dans la nanofeuille de graphène sont de nature différente des liaisons en violet dans le matériau SiC.

FIGURE 2

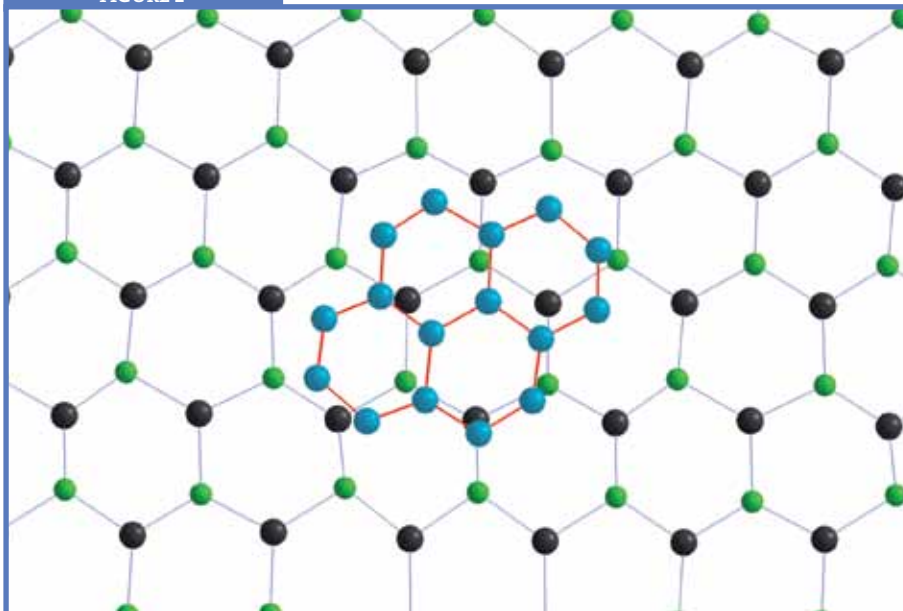
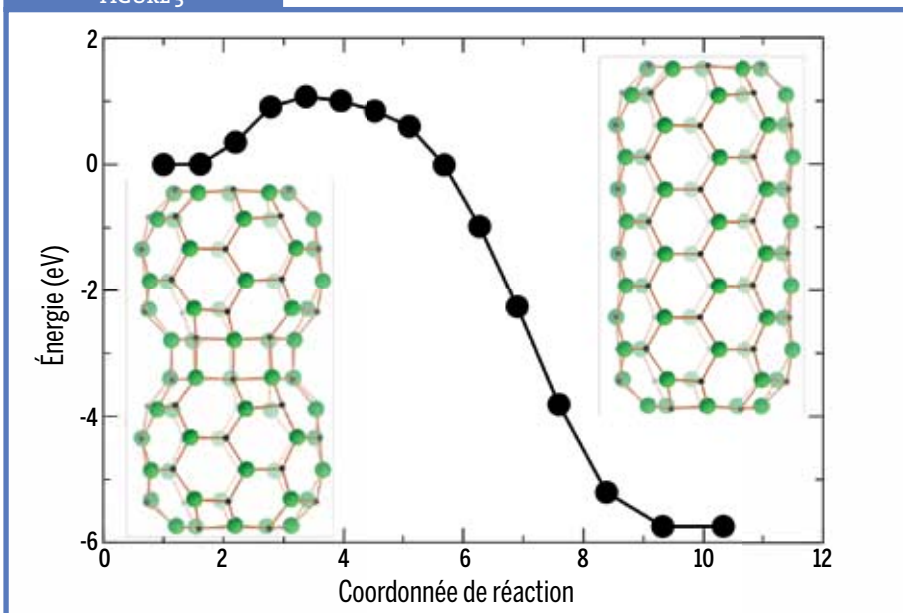


Figure 3

Courbe d'énergie représentant le chemin entre deux minima pour une cage de SiC (120 atomes). Il existe une barrière d'énergie (bosse) entre la structure de gauche (minimum local) et la structure de droite (minimum global).

Source : E. Machado-Charry et al., J. Chem. Phys., 135, 034102, 2011.

FIGURE 3



déterminer deux nouveaux *minima* par jour avec l'algorithme ART. Pour y parvenir, cet algorithme a été largement optimisé. Avec, à la clé, une diminution d'un facteur 4 du nombre d'évaluations d'énergies nécessaires pour déterminer un minimum et arriver à un ordre de 400 évaluations seulement – un travail préliminaire qui vient d'être publié dans le *Journal of Chemical Physics*.

Malgré tout, nous supposons qu'il nous faudra plusieurs mois afin d'obtenir des informations physiques exploitables dans la partie Monte Carlo cinétique.

Actuellement, nous étudions la croissance du graphène sur SiC en partant d'une surface terminée silicium (Figure 2). Ensuite, nous nous pencherons sur la croissance d'une feuille complète de graphène à partir des atomes de carbone présents dans la dernière couche du carbure de silicium. Avec l'espoir, à terme, de pouvoir confronter nos résultats aux données expérimentales. S'ils sont en bonne adéquation, nous aurons alors mis la main sur un mécanisme de synthèse du graphène sur carbure de silicium. ●

COMMENT CONCEVOIR UNE ARME ATOMIQUE SANS AVOIR À RÉALISER DE NOUVEAUX ESSAIS NUCLÉAIRES ? LA SOLUTION SE TROUVE CONJOINTEMENT DANS LA MODÉLISATION ET LA SIMULATION. DANS CES DOMAINES, LES SUPERCALCULATEURS ONT DÉJÀ FAIT LEURS PREUVES. DANS CET EXERCICE, LA FRANCE N'EST PAS EN RESTE.

LE BON CALCUL DE LA DISSUASION NUCLÉAIRE

Pour parvenir à simuler le fonctionnement complet d'une arme atomique sans avoir recours à de nouveaux essais nucléaires, il faut concevoir un modèle mathématique, résoudre un système d'équations sur des supercalculateurs et, enfin, valider les résultats obtenus grâce à des expériences menées en laboratoire et grâce aux mesures enregistrées lors des essais nucléaires du passé.

La première étape, celle de la modélisation, impose avant tout de bien connaître les différents phénomènes physiques impliqués et la façon dont ils s'enchaînent (lire « *Le b a-ba de la bombe H* »). Les équations susceptibles de les reproduire par le calcul sont connues : ce sont les équations de Navier-Stokes pour la mécanique des fluides, celles de Boltzmann pour le transport des neutrons et les équations de diffusion et de transport pour l'évolution de la matière et des photons. En couplant ces modèles physiques, on obtient un système d'équations mathématiques reproduisant fidèlement le fonctionnement d'une arme nucléaire.

Des milliards d'inconnues

Étape suivante, celle de la simulation pour résoudre ce système d'équations dans des conditions aussi proches que possible de la réalité du fonctionnement d'une arme nucléaire. Pour cela, il s'agit de décrire une grande variété de particules (des neutrons aux photons, en passant par les ions et les électrons) sur trois échelles de temps inférieures au millionième de seconde. Et dans des états extrêmes de pression : jusqu'à mille milliards de fois la pression atmosphérique ! Un calcul incroyablement plus complexe que beaucoup

d'autres modélisations, telle la météorologie, du fait de la brièveté des phénomènes en jeu et du couplage intime des mécanismes physiques.

Impossible de résoudre ce système de façon exacte. Il faut recourir à des méthodes d'analyse numérique pour le transformer en un système d'équations « approchées », soluble par un supercalculateur. Les calculs sont réalisés sur de petites zones, appelées des mailles. Plus les mailles sont nombreuses, plus on s'approche de la solution exacte du problème réel. Concrètement, des dizaines voire des centaines de millions de mailles sont utilisées. Cela représente des milliards d'inconnues.

Valider par partie

Pour s'assurer que l'ordinateur fournit une représentation aussi réaliste que possible, la validation des résultats se fait en deux phases. On valide d'abord des parties de modèle : un seul phénomène physique, tel le comportement mécanique d'un matériau, ou quelques phénomènes couplés. Pour cela, on compare les résultats des simulations à des expérimentations menées sur deux outils : la machine radiographique Airix et le Laser mégajoule (LMJ).

Installée à Moronvilliers en Champagne-Ardenne, depuis 2000, le Airix permet de radiographier une maquette d'arme nucléaire ne comportant pas de matière fissile pour valider la phase initiale, pyrotechnique. On y simule la compression des matières au moyen de matériaux non radioactifs aux comportements mécanique et thermique comparables. Les images radiographiques de cette phase de compression sont confrontées aux simulations numériques. Quant



PAR CHARLES LION
directeur du programme simulation à la direction des applications militaires du CEA.



LES ISOTOPES
sont des atomes qui ne se différencient que par leur nombre de neutrons.

LE FLOPS
(Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un téraflopp permet de faire mille milliards d'opérations par seconde (10^{12}).

au LMJ, en construction au Barp, près de Bordeaux, il devrait permettre à la fin 2014 de reproduire et d'étudier la phase de fusion nucléaire. Ce sera le pendant de la soufflerie pour un concepteur d'avion. Le LMJ pourra concentrer jusqu'à l'équivalent de 240 puissants faisceaux laser sur une bille remplie de deutérium et de tritium, deux isotopes (*) de l'hydrogène, provoquant leur fusion pendant quelques milliardièmes de seconde.

Un prototype, la Ligne d'intégration laser (LIL), comprenant quatre faisceaux a été mis en service en mars 2002 et a permis de valider les choix technologiques du LMJ. Celui-ci, comme c'est déjà le cas pour la LIL, sera ouvert à la communauté scientifique internationale – astrophysique, médecine, énergie, etc.

Arrive enfin la dernière étape : la validation globale. Elle consiste à comparer les résultats obtenus grâce à nos logiciels avec l'ensemble des mesures recueillies lors des essais nucléaires du passé, notam-

LE B A-BA DE LA BOMBE H

Première étape du fonctionnement d'une arme thermonucléaire : la détonation d'une charge explosive (pyrotechnie). En quelques milliardièmes de seconde, plusieurs milliers de degrés sont atteints, ce qui permet d'amorcer la fission en comprimant la matière fissile (plutonium ou uranium) transformée en plasma (gaz ionisé). Cette mini explosion nucléaire dure environ 100 milliardièmes de seconde et permet d'atteindre les dix millions de degrés requis pour amorcer la troisième étape, la fusion du deutérium et du tritium, deux isotopes de l'hydrogène. La température atteinte est alors de l'ordre du milliard de degrés pendant quelques milliardièmes de seconde.



Le Laser mégajoule (LMJ), en construction au Barp, près de Bordeaux, est un des outils majeurs pour la simulation des armes nucléaires. La fusion y sera reproduite en vraie grandeur dans cette chambre d'expérience de 10 mètres de diamètre.

juillet 2010, le supercalculateur est aujourd'hui pleinement opérationnel. Des simulations tridimensionnelles sont accessibles en grand nombre, ce qui permettra de garantir les prochaines têtes nucléaires des sous-marins qui devraient entrer en service en 2015. La simulation permettra d'adapter le standard de calcul à leurs environnements thermique et mécanique spécifiques. L'autre volet fondamental de ces recherches consiste à former et à homologuer les futurs concepteurs d'armes nucléaires. Ils doivent maîtriser parfaitement la simulation numérique et en connaître les limites.

La conjugaison des moyens uniques tels que les supercalculateurs Tera de classe mondiale et le LMJ permet aussi d'attirer de jeunes ingénieurs, physiciens et mathématiciens brillants et de conserver les compétences nécessaires à nos missions de dissuasion. ●

ment ceux effectués en 1996, lors de la dernière campagne. Cette étape permet de produire un standard de calcul, autrement dit de véritables prescriptions d'emploi des simulations numériques, le domaine de confiance de la simulation à un temps «t».

Une première mondiale

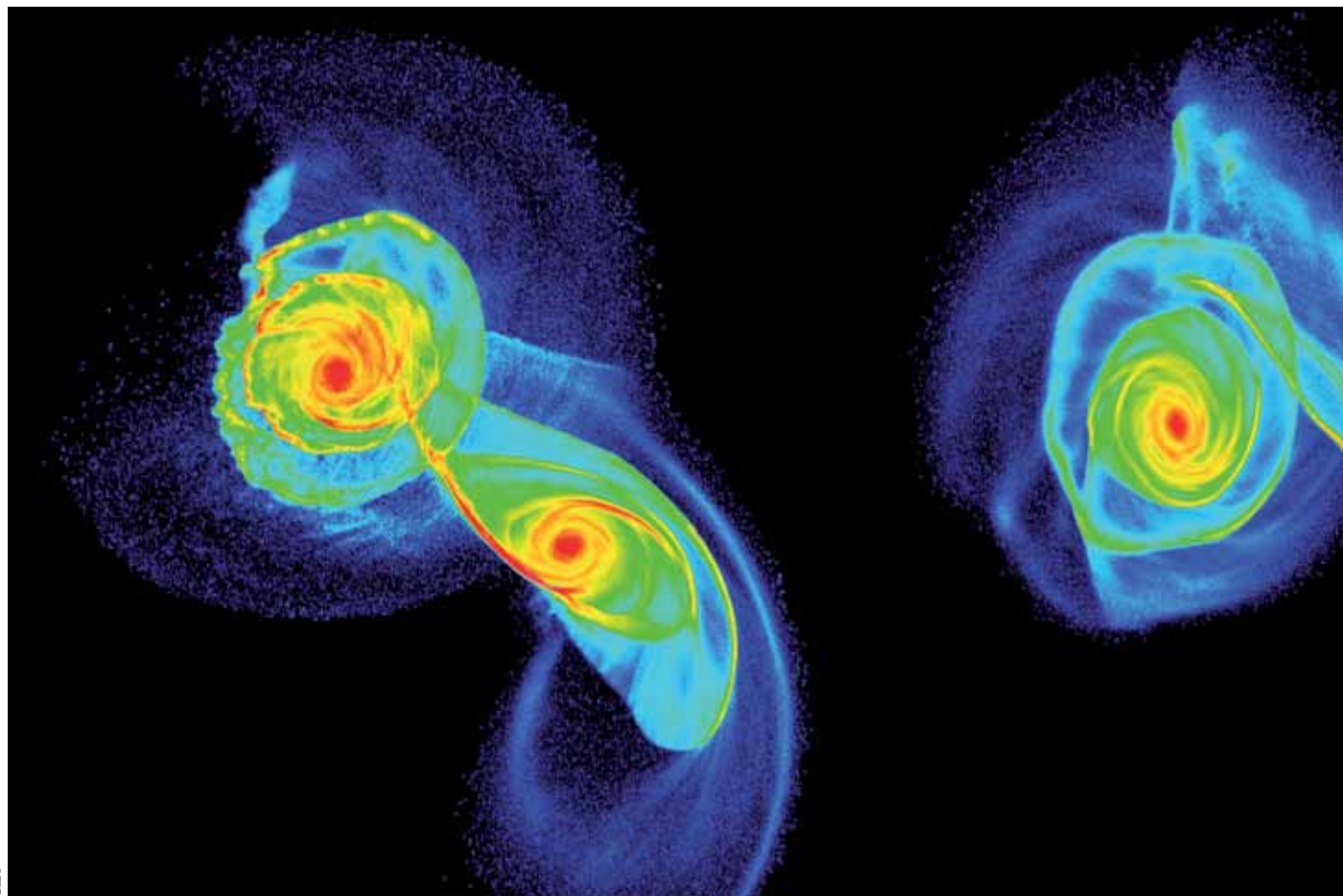
En 2001, grâce au supercalculateur Tera 1 d'une puissance de calcul de 5 téraflops^(*), installé au CEA de Bruyère-le-Chatel (Essonne), nous avons produit notre premier standard de calcul, premier pas pour garantir les armes nucléaires par la seule simulation. En 2005,

le Tera 10 (50 téraflops) a permis de définir un standard de calcul validé sur un ensemble d'expériences bien plus large, avec les premières simulations tridimensionnelles. Nous avons pu garantir le fonctionnement d'une tête nucléaire aéroportée par la seule simulation sans nouvel essai nucléaire : une première mondiale à notre connaissance. Les missiles portés par le Mirage et le Rafale ont commencé à en être équipés dès 2009.

Avec une puissance de 1 000 téraflops, le Tera 100 permet un nouveau saut significatif dans les capacités de modélisation. Installé en

QUE SE PASSE-T-IL LORSQUE DEUX GALAXIES VIENNENT À SE PERCUTER ? VOILÀ UNE QUESTION QUI NE RELÈVE PAS QUE DE LA SIMPLE CURIOSITÉ DES ASTRONOMES : CAR LES PROCESSUS D'INTERACTIONS ET DE FUSIONS DES GALAXIES SONT DES PHASES CLÉS DANS LA NAISSANCE ET LA FORMATION DES ASTRES.

COMPRENDRE COMMENT SE FORMENT LES ÉTOILES



CEA

Au début de l'Univers commencent à se former de petites galaxies naines, qui vont fusionner entre elles pour former des galaxies de plus en plus massives. Les fusions de galaxies participent aussi à la redistribution du moment angulaire^(*) entre la composante visible des galaxies en rotation rapide (gaz et étoiles) et la composante de matière noire, laquelle s'étend comme un halo autour de la matière visible. Elles favorisent la chute de gaz au centre des galaxies,



PAR PAOLA DI MATTEO, FRANÇOISE COMBES ET BENOÎT SEMELIN

astrophysiciens à l'observatoire de Paris.

où d'intenses flambées de formation stellaire peuvent se produire.

Un mois de simulation

Voilà pour la théorie. Comment maintenant décrire ces phénomènes avec précision ? Faute de pouvoir les mesurer, sachant qu'ils se produisent à des distances astronomiques de nous et sur des échelles de temps dépassant largement la vie humaine... En simulant l'évolution de galaxies en interaction ! Facile à énoncer, ce

problème est un véritable défi scientifique. Pour simuler le comportement des galaxies, il faut prendre en compte la dynamique de leurs disques à l'échelle de dizaines de kiloparsecs (kpc)^(*), soit des distances de l'ordre de 10^{21} m (mille milliards de milliards de mètres). Quant à la formation des étoiles et des nuages interstellaires, l'échelle à prendre en compte est de l'ordre du parsec, soit environ 10^{16} m. Et les calculs doivent en outre coupler les deux échelles. Les supercalculateurs nous ont

permis d'amorcer une étude statistique des fusions de galaxies, en réalisant une cinquantaine de simulations, à une résolution spatiale de l'ordre de 50 parsecs. Concrètement, nous avons mené des calculs pendant plus d'un mois sur environ 1 550 cœurs du supercalculateur Curie, installé au Centre de calcul, recherche et technologie du CEA (CCRT-CEA).

Modélisation des fusions

Nous modélisons ainsi la fusion de galaxies de même taille – on parle de « fusions majeures » – ainsi que l'accrétion de satellites sur une galaxie du type Voie lactée – ou « fusions mineures ». Les résultats de cette évolution violente



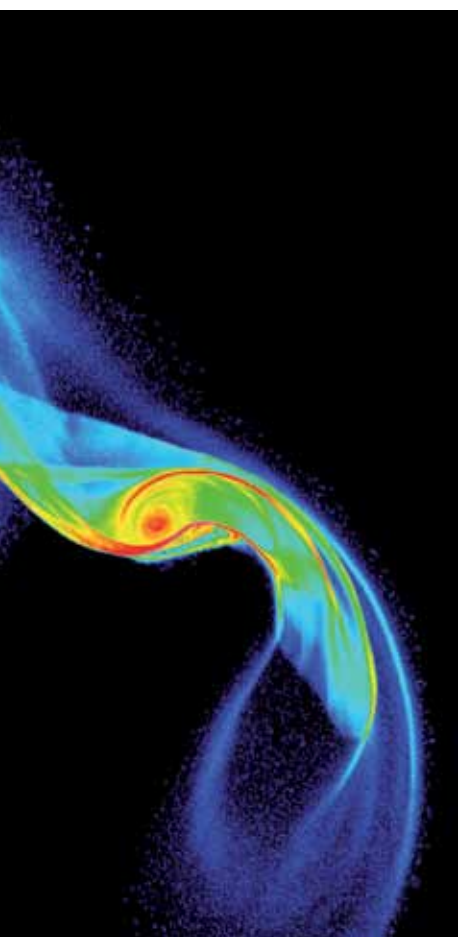
MOMENT ANGULAIRE

Vecteur parallèle à l'axe de rotation dont l'amplitude représente la quantité de rotation d'un système.

PARSEC

Parallaxe-seconde, unité de longueur astronomique, équivalent de 3,26 années-lumière, ou 206 265 unités astronomiques (distance Terre-Soleil)

Distribution du gaz après la première rencontre de deux galaxies de même taille (fusion majeure). Lors du passage proche de deux galaxies, une partie de la matière peut être expulsée à de larges distances du disque, avec la formation de structures dites « queues de marée », qui peuvent s'étendre sur des centaines de kiloparsecs.



peuvent ensuite être comparés à l'évolution des processus internes plus lents, dits d'évolution séculaire. L'objectif final est de comprendre notamment la redistribution du moment angulaire entre la matière visible et la matière noire, l'intensité des sursauts de formation stellaire qui se produisent, ou encore les propriétés morphologiques des résidus de fusions.

Dans notre méthode de simulation, le système est modélisé par un ensemble de « particules » dont

on suit l'évolution sous l'effet des différents processus physiques pris en compte. Une particule correspond soit à un amas d'étoiles, soit à un nuage de gaz interstellaire. Les processus physiques sont essentiellement les forces de gravité qui s'exercent entre toutes les particules, mais aussi les forces de pression du gaz, la viscosité, les ondes de choc, la formation des étoiles à partir du gaz, et l'éjection de gaz par les étoiles en fin de vie.

Pour obtenir une résolution spatiale de 50 parsecs, la masse des particules doit être de l'ordre de 5 000 fois celle du Soleil, et il faut 30 millions de particules pour modéliser la fusion de deux galaxies du type de la Voie lactée. Pour évaluer les forces gravitationnelles entre N particules, les algorithmes les plus simples demandent N^2 calculs d'interactions. Au prix d'une approximation, l'algorithme auquel nous avons fait appel, dit « en arbre » et développé dans les années 1980, permet de réaliser le calcul en un ordre d'opération proportionnel à $N \ln(N)$ opérations : sans ce gain substantiel, ces simulations seraient impossibles.

Forces hydrodynamiques

Mais les forces gravitationnelles ne sont pas tout. Il faut aussi prendre en compte les forces hydrodynamiques (pression et viscosité) caractérisant la dynamique des galaxies. Nous les calculons par la méthode dite SPH (Smooth Particle Hydrodynamics), développée à partir de la fin des années 1970, qui consiste à représenter le fluide en une multitude de petits éléments se recouvrant.

Pour être capable de suivre non seulement la rotation des galaxies mais aussi la formation lors des fusions de petites structures denses, de type nuages moléculaires dont le temps d'évolution dynamique est beaucoup plus court, nous évaluons les forces tous les 250 000 ans, c'est-à-dire 8 000 fois sur les 2 milliards d'années qu'englobe une simulation. On peut estimer que cela correspond au total à environ 10 000 milliards de calculs d'interactions entre particules pour une seule simulation.

Si nous n'avions qu'un seul processeur à disposition, il faudrait plusieurs années de calcul. Seule solution : faire appel à plusieurs unités de calcul pour chaque simulation. Dans le cadre du « grand Challenge » sur le supercalculateur Curie, et grâce à la paralléli-

sation du code (OpenMP), chaque simulation a été lancée sur 32 cœurs et a utilisé plus de 50 Go de RAM sur chaque nœud de calcul. La force de ces simulations – toujours en cours et dont l'exploitation commence à peine – est de nous apporter des résultats avec une résolution spatiale inédite. Dans un premier temps, celle-ci nous permettra de savoir où et avec quelle efficacité le gaz est transformé en étoiles.

Notre vision de la dynamique des galaxies en sera radicalement changée. En effet, l'énergie nucléaire dans les étoiles est non seulement « rayonnée », mais aussi transformée en énergie cinétique, au travers des flots de gaz rejetés. À basse résolution, la formation d'étoiles ne peut être traitée que de façon semi-analytique, en se servant d'une probabilité proportionnelle à la densité de gaz présente. Or, cela n'est pas satisfaisant car les instabilités des disques des galaxies sont plus subtiles.

La physique globale fragilisée ?

L'espoir est donc de traiter le gaz interstellaire de façon plus fine, en considérant sa dissipation par rayonnement, les collisions entre nuages, les ondes de choc formées, la turbulence macroscopique générée et sa viscosité. En prenant en compte ces différents paramètres physiques avec précision, il n'est pas dit que la physique globale n'en soit pas complètement transformée. D'abord parce que l'énergie réinjectée dans le milieu interstellaire par les étoiles, les vents stellaires et explosions de supernovae, peut passer inaperçue, si l'échelle considérée est inférieure à la résolution. En revanche, si les simulations ne l'ignorent pas, alors cette énergie transférée dans le milieu interstellaire change complètement sa dynamique et peut empêcher la formation d'étoiles.

Mais surtout, le milieu interstellaire comporte plusieurs phases, dont les densités sont très différentes, de plus de 10 ordres de grandeur entre les nuages les plus denses et le milieu diffus. À basse résolution, les diverses phases peuvent être simulées par des composants différents, avec des échanges de masse entre phases, calibrés de façon semi-analytique. À haute résolution, on peut espérer recréer directement plusieurs phases, par l'intermédiaire des instabilités naturelles du milieu, et ainsi traiter les nuages moléculaires à part entière, sans les introduire de façon artificielle. ●

LA MÉCANIQUE QUI AFFECTE UN MATÉRIAU EN CAS DE CHOC INTENSE RESTE MAL COMPRISE. POUR L'Étudier, il faut pour cela descendre au niveau de l'atome. DÉFORMATION IRRÉVERSIBLE, ENDOMMAGEMENT, DÉCOMPOSITION... LA SIMULATION DE HAUT VOL N'EST PAS DE TROP POUR MIEUX COMPRENDRE LA PHYSIQUE DES CHOCS.

LA PHYSIQUE DES CHOCS À L'ÉCHELLE ATOMIQUE

Que se passe-t-il au niveau moléculaire lorsqu'un matériau subit un choc intense consécutif, par exemple, à une explosion ou un violent impact ? Le choc peut être vu comme une compression extrêmement rapide – de l'ordre de 10^{-11} s – et intense – 10^{10} Pascal, soit 10^5 atmosphères et plus... Il est encore impossible d'observer expérimentalement avec précision et en temps réel ce qui se passe à l'échelle de l'atome. On sait néanmoins qu'il s'agit de phénomènes complexes : plasticité (déformation irréversible du matériau), endommagement (fissuration, rupture), décomposition chimique, etc. Pour avoir une idée plus précise, une solution passe par des simulations sur des ordinateurs pétaflopiques.

La dynamique moléculaire est une méthode de simulation largement utilisée. Elle consiste à résoudre les équations du mouvement d'un ensemble de particules (atomes, molécules) interagissant par le biais d'une force prédéfinie. Ces dernières années, les calculateurs permettaient de réaliser des études plutôt phénoménologiques sur des systèmes de dimension réduite – quelques millions d'atomes soit des dimensions de quelques millièmes de micron.

Avec la machine Tera 100, on peut désormais prendre en compte des systèmes dont les dimensions, proches du micron, permettent une comparaison avec l'expérience, tout en utilisant des forces d'interaction décrivant avec une bonne approximation la complexité des interactions interatomiques.

Pour mener à bien ces simulations, il nous a fallu repenser les logiciels afin qu'ils tirent pleinement profit des capacités des cal-



PAR LAURENT SOULARD, NICOLAS PINEAU, OLIVIER DURAND ET JEAN-BERNARD MAILLET
ingénieurs à la Direction des applications militaires du CEA.

culateurs massivement parallèles : c'est le cas du code Stamp, développé à la Direction des applications militaires du CEA (CEA-DAM) depuis une dizaine d'années. Forts de ce code, nous avons pu simuler quelques cas très concrets.

Premier exemple, la transition d'un régime élastique vers un régime plastique : au-delà d'un certain seuil de pression, la propagation d'un choc entraîne la création irréversible de défauts structuraux. Parce qu'elle affecte considérablement les propriétés mécaniques du matériau, la plasticité nécessite de procéder à une modélisation très fine en s'appuyant sur une description des mécanismes élémentaires sous-jacents.

Plasticités et rupture

Nous nous sommes intéressés en particulier à la plasticité du diamant : un véritable défi car elle impose non seulement de prendre en compte la complexité du processus, mais aussi l'implantation d'un modèle d'interaction capable de reproduire les nombreux états du carbone (diamant, graphite, etc.). Le code Stamp permet de décrire très précisément le carbone, mais au prix d'un temps de calcul important. Nous l'avons utilisé pour simuler la propagation de chocs d'intensités diverses sur les principales orientations cristallographiques du diamant (figure 1).

Ainsi, pas moins de 1,3 million d'atomes de carbone ont été pris en compte pour le calcul effectué par 6 400 cœurs de la machine Tera 100. Résultat : nous avons vu apparaître de multiples défauts structuraux traduisant l'apparition de la plasticité. Cette mesure très détaillée de l'état de la matière sera une aide précieuse pour la construction d'un modèle utili-

sable dans un code travaillant à l'échelle macroscopique.

Un choc rencontre au cours de sa propagation des interfaces entre matériaux, et en particulier des surfaces dites « libres », formant les limites entre le système et l'air ambiant. Or, lorsqu'il se réfléchit sur une surface libre, le choc peut être à l'origine de différents phénomènes qui dépendent de son intensité ainsi que de la topologie locale de la surface – entre autres sa rugosité. Ces phénomènes peuvent fortement affecter les propriétés du matériau et endommager son environnement comme, par exemple, un système de mesure.

C'est notamment le cas de la « rupture par écaillage ». Ce phénomène se produit lorsque l'onde de détente (ou décompression) issue de la réflexion du choc sur la surface libre rencontre celle qui suit le choc initial. Il se produit localement une mise en tension très brutale du milieu qui peut entraîner la rupture du matériau. L'écaillage fait activement l'objet de travaux associant des simulations de dynamique moléculaire à grande échelle et des expériences spécifiques.

Éjection de matière

Autre phénomène induit par le choc : une projection de matière. Celle-ci est provoquée par un défaut de planéité de la surface libre, comme par exemple une rayure créée lors du passage d'un outil d'usinage. Selon le défaut, la matière éjectée se présente sous la forme de petits agrégats, de jets, etc. Si l'on veut protéger les matériaux environnant (appareil de mesure, revêtement de la chambre d'expérience, etc.), il faut maîtriser ces mécanismes d'éjection. Les conditions de formation d'un jet, par exemple, obéissent à des conditions hydrodynamiques relative-

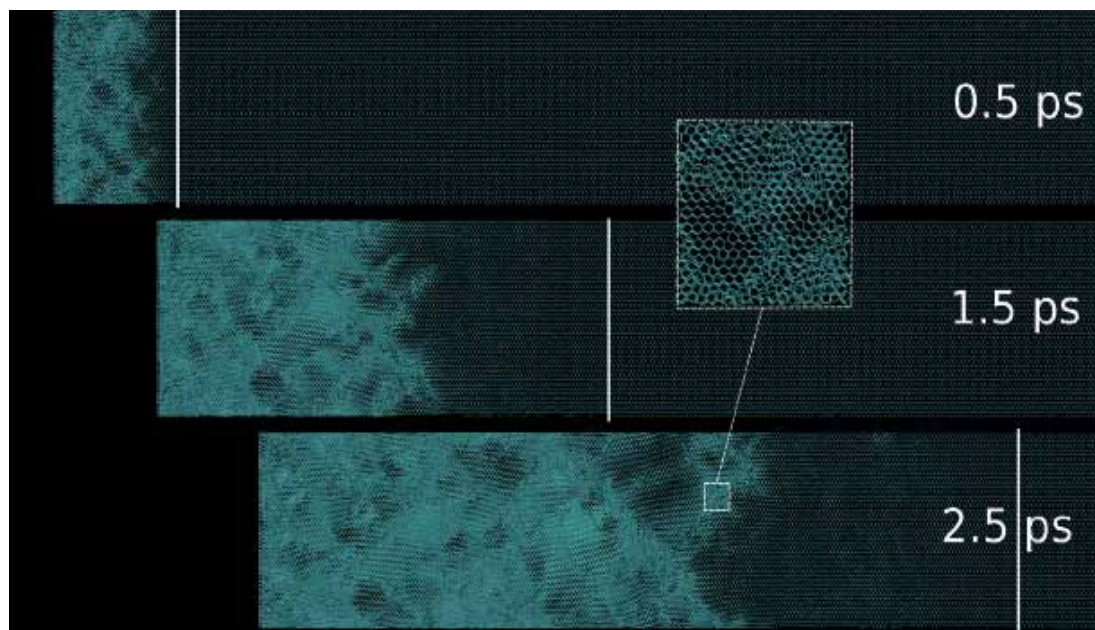


Figure 1. Propagation d'une onde de choc dans un monocristal de diamant derrière le front de choc (trait blanc), de multiples défauts structuraux traduisent l'apparition de la plasticité (1 ps vaut 1 picoseconde, soit 10^{-12} seconde).

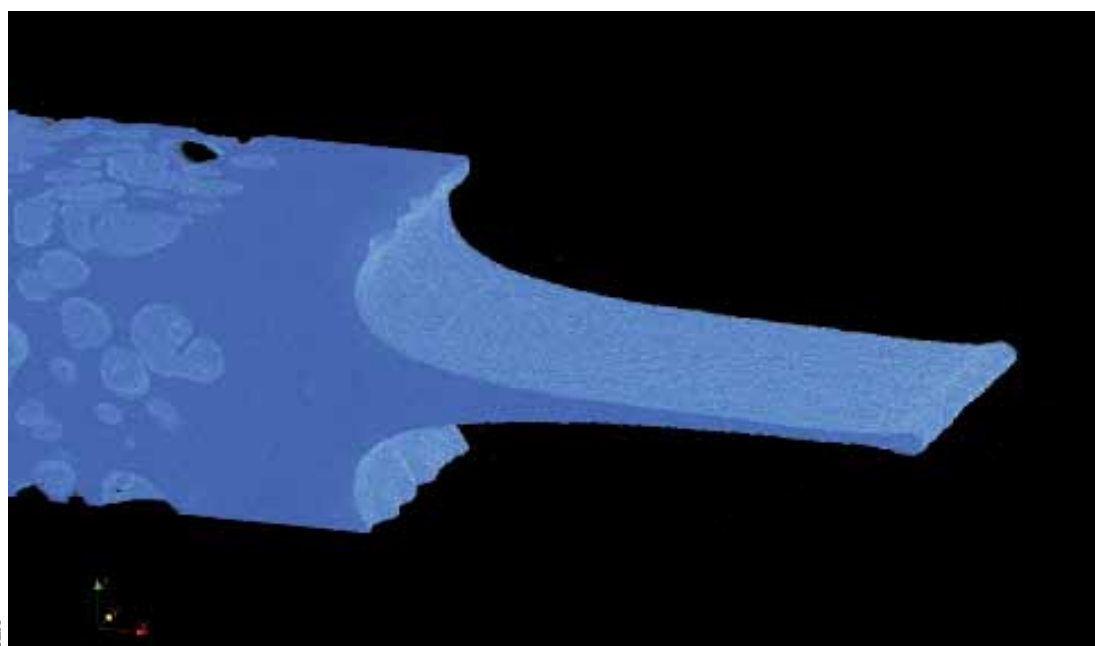


Figure 2. Formation d'un jet de matière lors de la réflexion d'un choc sur une surface affectée d'un défaut de planéité. Un total de 125 millions d'atomes de cuivre et 4 000 cœurs du Tera 100 ont été requis pour ce calcul.

CEA

ment bien connues. Son comportement ultérieur, et notamment sa fragmentation, est en revanche beaucoup plus difficile à appréhender. Là aussi, on peut aborder le problème de façon extrêmement fine – mais sur une échelle spatiale réduite – grâce à une expérience numérique de dynamique moléculaire. Nous avons ainsi simulé la formation d'un jet en prenant en compte 125 millions d'atomes de cuivre grâce aux calculs menés par 4 000 cœurs sur la machine Tera 100 (figure 2). Et comme avec l'étude de la plasticité, ces informations nous permettront de construire des modèles aux échelles macroscopiques.

Enfin, si le matériau dans lequel se propage le choc est le siège de réactions chimiques, on peut at-

teindre un régime de détonation, c'est-à-dire la propagation d'une réaction chimique par onde de choc. On ne peut bien comprendre ce phénomène qu'en étudiant les mécanismes moléculaires qui en sont à l'origine. La décomposition chimique de l'explosif se produit sur une épaisseur allant de quelques microns à quelques millimètres, selon la nature de l'explosif. Or, les simulations de réactions chimiques, qui nécessitent par nature de prendre en compte des électrons, sont limitées aux échelles nanométriques.

Revisiter les modèles

Comment lier les deux échelles ? Par un modèle qui permet de simuler l'évolution d'un ensemble de « superparticules » représentant

une ou plusieurs molécules possédant leur propre dynamique interne. La dynamique de ce système est donnée par une extension de la dynamique hamiltonienne usuelle ; elle reste donc compatible avec un code de dynamique moléculaire. Des simulations du nitrométhane (CH_3NO_2) avec 8 000 cœurs sur le Tera 100 ont permis de montrer que l'apparition d'une détonation est un processus beaucoup plus complexe que le laisse supposer l'observation expérimentale.

La simulation permet de revisiter les modèles macroscopiques. Elle complète les expériences pour comprendre cette physique complexe. Étroitement liée à la puissance des ordinateurs, son importance ne pourra que croître dans les prochaines années. ●

DEPUIS ENVIRON QUARANTE ANS, DES ALLIAGES MÉTALLIQUES RÉVÈLENT DES PROPRIÉTÉS MÉCANIQUES SURPRENANTES : ILS ENCAISSENT DE GRANDES DÉFORMATIONS AVANT DE RETROUVER LEUR FORME INITIALE. LA COMPRÉHENSION DE CE PHÉNOMÈNE AU NIVEAU ATOMIQUE NE POURRAIT S'ENVISAGER SANS DES PUISSANCES DE CALCUL COLOSSALES.

LES DÉFORMATIONS MARTENSITIQUES SOUS LA LOUPE DES PROCESSEURS

C'est en 1970 que l'on a découvert des alliages métalliques aux propriétés surprenantes : sous une faible contrainte de traction, ils peuvent s'étirer d'un certain pourcentage, puis retrouvent, quand on les «relâche», leur forme initiale. Ces alliages complexes – souvent à base de nickel et de titane –, dont le comportement est quelquefois appelé «pseudo-élastique», ont rapidement trouvé de nombreuses applications, comme les montures de lunettes ou les stents, en chirurgie.

Le mécanisme permettant d'aussi grandes déformations est appelé «transformation martensitique» (fig 1) : il s'agit d'une modification de la maille cristalline consécutive à une modification de l'environnement, soit une contrainte appliquée, une pression, ou un changement de la température. Dans les métaux, ces transformations sont extrêmement courantes et elles ont fait l'objet d'une intense activité de recherche. Bien qu'elles permettent rarement des comportements aussi spectaculaires que la pseudo-élasticité, elles peuvent induire, sous l'action de fortes sollicitations, d'importantes déformations, et jouent un rôle fondamental dans le comportement des métaux.

Notre compréhension de ces mécanismes de déformation à l'échelle atomique a beaucoup progressé grâce à des modélisations de transitions martensitiques sous choc par dynamique moléculaire, réalisées sur des millions d'atomes. Pour des sollicitations prolongées, la microstructure finale devient très complexe et s'étend sur de larges zones d'espace, souvent de l'ordre

du millimètre. Peut-on simuler les transitions martensitiques en dynamique moléculaire sur d'aussi larges domaines ?

Mur du temps

Lors d'une telle simulation, les équilibres s'établissent via la propagation d'ondes de déformation. La distance maximale parcourue par ces ondes, proportionnelle au temps de la simulation, limite la corrélation entre différents points de l'espace. En pratique, augmenter la taille de la boîte de calcul au-delà de cette «distance maximale de corrélation» n'apporte rien de plus si l'on n'augmente pas la durée de la simulation dans les mêmes proportions. Or, si la puissance des calculateurs augmente grâce à la multiplication des cœurs, la puissance par cœur a en revanche peu évolué.

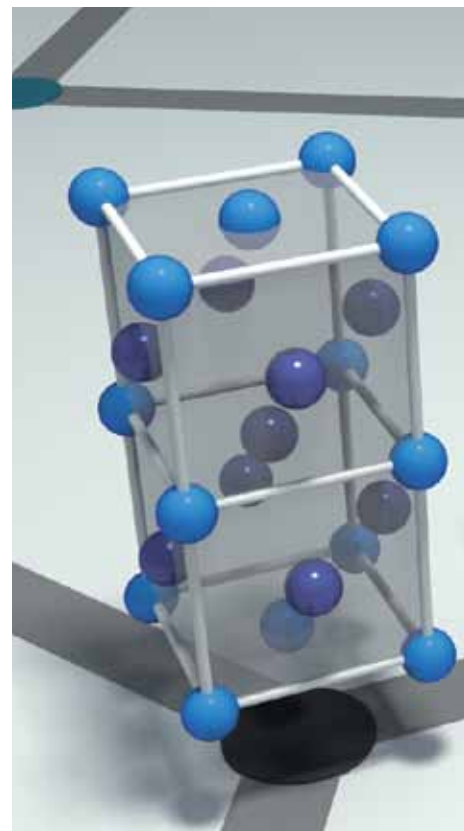
Aujourd'hui, comme il y a dix ans, il est en pratique difficile de dépasser le million de cycles de calculs, chacun d'eux correspondant à un pas de temps d'une femtoseconde (un millionième de milliardième de seconde), soit une durée totale simulée d'une nanoseconde. Cette durée limitée correspond à une distance maximale de corrélation de quelques micromètres seulement et restreint l'apparition de microstructures de tailles supérieures.

Les plus grandes simulations se heurtent donc à ce «mur du temps». On ne peut le franchir qu'en changeant de représentation, par exemple en regroupant quelques atomes d'une molécule complexe pour en optimiser le calcul. Dans le même esprit, nous avons opté en 2010 pour une méthode qui permet un changement radical de la «granularité» de la représentation de la matière en regroupant les atomes par 10, 100



PAR
CHRISTOPHE
DENOUIL

ingénieur
de recherche
au CEA.



ou 1 000, tout en assurant une certaine équivalence entre l'énergie de cet ensemble et la représentation atomique (et détaillée) sous-jacente.

Cette approche repose sur une représentation particulièrement compacte du paysage énergétique impliqué dans les changements de phases. Chacun des états stables est représenté par un minimum d'énergie, dont le niveau est estimé (avant le calcul) par une technique de modélisation à l'échelle atomique. Lors d'une transition, la matière passe d'un état stable (ou puits) à un autre en empruntant un «chemin de réaction», c'est-à-dire un ensemble d'états qui permettent

• Transformations martensitiques : les alliages de fer-nickel peuvent facilement passer d'une structure cubique centrée (en haut à gauche) à cubique centrée (en haut à droite), moyennant une déformation homogène de la maille (cadre blanc). Les transitions de phases sont représentées comme des lignes (en gris) connectant deux états stables (pastilles de couleur). La dimension de l'espace de l'arbre (ici représenté schématiquement en deux dimensions) est dans le cas général un espace à 9 dimensions.

une transition facile entre deux puits. Cet arbre des chemins de réaction (figure 1) est dupliqué pour toutes les mailles du calcul, chacune minimisant au mieux l'énergie.

L'arbre des chemins de réaction

Prenons l'exemple du calcul sur un alliage de fer et de nickel (figure 2). Réalisé avec le Tera 100 sur plus de 4500 processeurs, il a permis d'atteindre des dimensions de boîtes de 0,5 micron^(*) de côté pour un temps de simulation de 1 microseconde. Les microstructures qui émergent de ces calculs font apparaître une alternance de lamelles composées de différentes variantes de martensites aux fron-

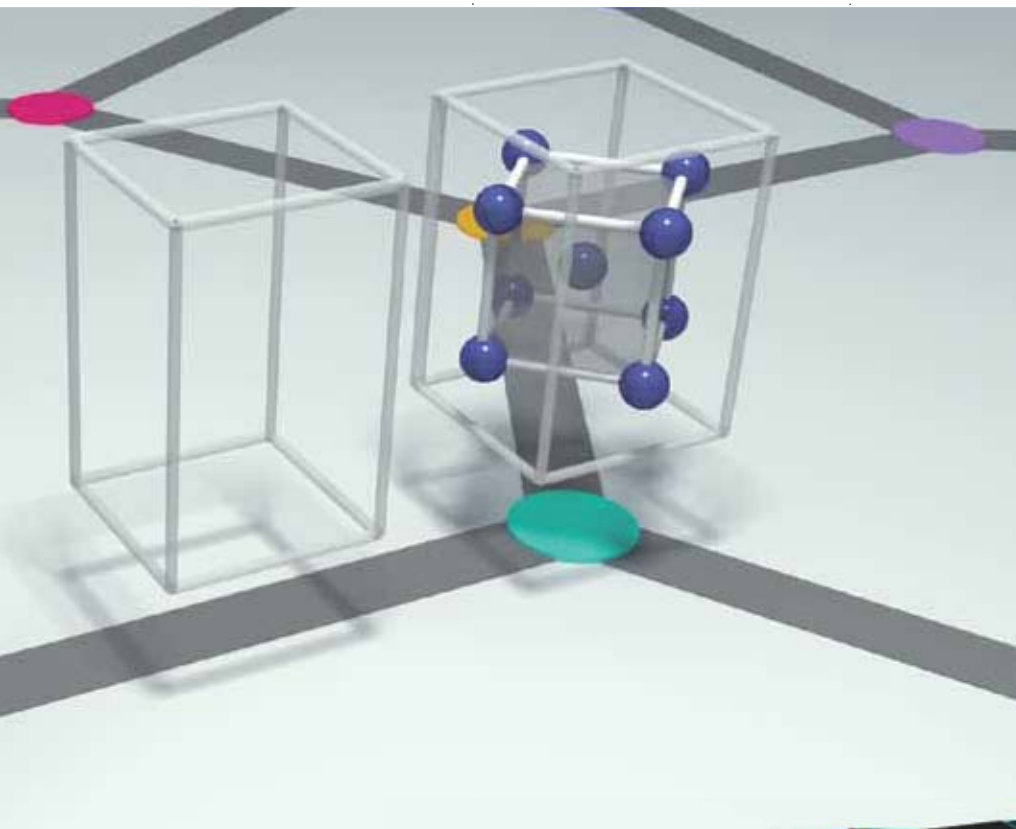

LE MICRON,
ou micromètre,
vaut un millième
de millimètre (10⁻⁶m).

tières bien planes. À plus grande échelle, elles s'empilent également en bandes relativement rectilignes, elles-mêmes contenues dans un couloir plus large.

Pour pouvoir obtenir cette structure imbriquée sur trois niveaux, il faut pouvoir garantir à la fois une très bonne résolution – ici chaque maille représente une centaine d'atomes –, de grandes boîtes de calcul et un temps de calculs suffisamment long pour que de grandes structures apparaissent. Seuls des calculs à grands grains, réalisés sur plusieurs milliers de cœurs permettent d'atteindre à la fois ces résolutions et ces échelles d'espace et de temps.

Les résultats que nous obtenons aujourd'hui reposent sur un arbre de chemins de réaction calculé à l'aide de potentiels atomiques simplifiés. Mais en réalité, ces transformations sont induites par une recomposition de la structure électronique des atomes. Or, à l'échelle de l'atome, ces transformations sont décrites uniquement par une approche quantique. Les simuler demande de recourir à une autre approche : les calculs de structures électroniques dits *ab initio*. Ils sont également très complexes et très consommateurs de temps de calculs et bénéficient eux aussi du parallélisme massif, propre au calcul haute performance. On obtient ainsi une estimation fiable des chemins de réaction.

En couplant ces deux approches, nous aurons bientôt une vision unifiée d'une microstructure de l'échelle du millimètre à l'échelle de l'atome. Ces transformations, quelquefois si rapides qu'elles échappent aux diagnostics les plus fins, pourront alors être analysées dans le détail, et le mystère de leur formation partiellement levé. •



• L'apparition des lamelles : une déformation, appliquée ici en 1 microseconde sur un alliage de fer et de nickel, induit une transformation martensitique complexe. Les différentes phases s'empilent en grandes lamelles et se structurent elles-mêmes en bandes.



CEA

LES PREMIERS RAYONS LASER ONT ÉTÉ PRODUITS IL Y A PLUS DE CINQUANTE ANS ! MAIS LE COMPORTEMENT DE CETTE LUMIÈRE LORSQU'ELLE TRAVERSE DES MATÉRIAUX RESTE LA SOURCE DE NOMBREUSES INTERROGATIONS. OPTIQUE NON-LINÉAIRE, AUTOFOCALISATION... GRÂCE AUX SUPERCALCULATEURS, LE RAYON LASER PEUT ENCORE NOUS SURPRENDRE.

DES PROCESSEURS GRAPHIQUES POUR VISUALISER LA LUMIÈRE

C'est un des enseignements de base de l'optique dite « linéaire » : la lumière se comporte différemment selon le milieu qu'elle traverse. À chaque matériau son indice de réfraction. Mais c'est une vérité qui a été mise à mal dans les années 60 avec l'avènement des sources laser : l'indice de réfraction des matériaux transparents – les gaz, les verres... – peut dépendre de l'intensité lumineuse qui les traverse. L'étude de ce phénomène s'appelle « l'optique non linéaire ».

Sous certaines conditions, par exemple lorsque la puissance délivrée par un laser devient supérieure à une valeur seuil, l'indice du milieu augmente continuellement le long du chemin optique. Conséquence : l'impulsion laser se focalise à la manière d'une loupe ou d'une lentille. On parle alors d'« autofocalisation optique ».

Architecture obsolète

Cette étrange propriété de la lumière se rencontre dans des expériences utilisant des sources laser d'énergie modérée – quelques millijoules – mais délivrant des impulsions de durée femtoseconde^(*). Elle se rencontre aussi si l'on a affaire à des sources de forte énergie (une dizaine de kilojoules) mais pour des impulsions de durée, plus longue, nanoseconde^(*). C'est le cas par exemple avec les installations laser de forte puissance, où le phénomène d'autofocalisation se manifeste dans les verres de silice et conduit à une fragmentation de l'impulsion optique en une multitude de filaments de taille micrométrique et d'intensité mille fois supérieure à celle de l'onde incidente.

Pour décrire ces dynamiques non linéaires, on réalise des simulations numériques avec des super-



PAR
LUC BERGÉ

directeur de recherche
au département
de physique
théorique
et appliquée,
au centre CEA
de Bruyères-
le-Châtel (91)

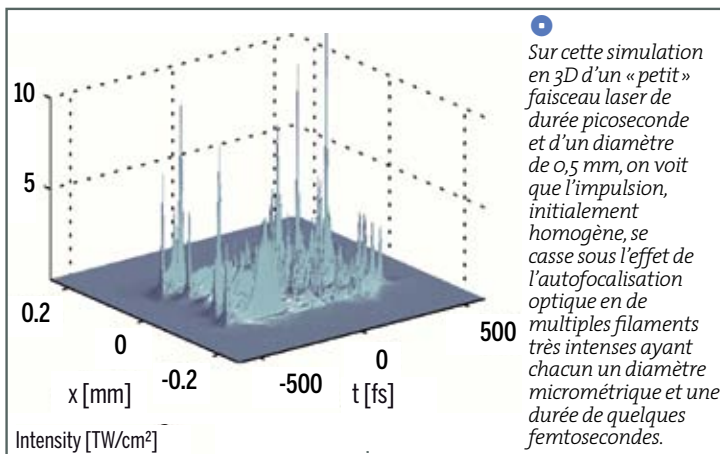
ET GUILLAUME
COLIN DE
VERDIÈRE

ingénieur-
chercheur
au département
des sciences de
la simulation et
de l'information,
au centre CEA
de Bruyères-
le-Châtel (91).



LA FEMTOSECONDE

(1 fs) vaut 10^{-15} s.
La nanoseconde
(1 ns) vaut 10^{-9} s.
La picoseconde
(1 ps) vaut 10^{-12} s.
L'intensité d'une
onde lumineuse
s'exprime en
watt par cm^2 .
Un térawatt (TW)
vaut 10^{12} W.



Sur cette simulation en 3D d'un « petit » faisceau laser de durée picoseconde et d'un diamètre de 0,5 mm, on voit que l'impulsion, initialement homogène, se casse sous l'effet de l'autofocalisation optique en de multiples filaments très intenses ayant chacun un diamètre micrométrique et une durée de quelques femtosecondes.

calculateurs à haute performance, comme les machines Titane ou Tera 100, situées au centre CEA de Bruyères-le-Châtel (Essonne). Mais les techniques de parallélisation classiques, qui consistent à calculer le champ laser simultanément sur différents processeurs CPU (« Central Processing Unit ») distribués le long d'une seule dimension spécifique les uns à la suite des autres, deviennent obsolètes pour ce type de calcul.

Décrire la formation de plasma

En effet, pour simuler une impulsion laser de diamètre inférieur au millimètre, il faudrait monopoliser 128 processeurs CPU pendant plusieurs mois. Avec des faisceaux de taille centimétrique tels que ceux du Laser Megajoule (LMJ), tout doit être multiplié par 10 000 ! De plus, outre les mécanismes optiques, il faut aussi simuler la dégradation du matériau. Concrètement, cela conduit à calculer la génération de plasma, produite pendant quelques femtosecondes, lorsque les intensités laser atteignent plusieurs dizaines de TW/cm^2 (*).

La solution passe par l'utilisation d'un autre type de processeurs :

les processeurs graphiques, ou GPU pour « Graphics Processing Units ». Originellement destinés aux jeux vidéo, les GPU récents possèdent une puissance de traitement extrêmement importante grâce à plusieurs centaines d'unités de calcul pouvant travailler en parallèle.

L'art du programmeur consiste à organiser toutes ces ressources en exprimant son algorithme à l'aide de milliers de processus légers, selon un découpage particulier dit « grille de calcul ». Les accélérations de calcul entre processeurs GPU et leurs ancêtres CPU peuvent atteindre un facteur 50 et résoudre en quelques heures des systèmes d'équations non linéaires qui nécessitaient plusieurs jours d'intégration numérique il y a un an.

Actuellement, les calculs sur GPU offrent la possibilité de simuler des impulsions laser de 3 nanosecondes avec une résolution de 30 femtosecondes (soit cinq ordres de grandeur sur une dimension) en moins d'une semaine. Prochain défi : atteindre la femtoseconde pour décrire la formation de plasma. Des processeurs graphiques pour visualiser la lumière, c'est bien naturel... ●

L'ACCROISSEMENT DE LA PUISSANCE DES SUPERCALCULATEURS PASSE PAR UNE AMÉLIORATION DRASTIQUE DE LEUR EFFICACITÉ ÉNERGÉTIQUE. MÉMOIRES ET PROCESSEURS MOINS GOURMANDS, ARCHITECTURE MASSIVEMENT PARALLÈLE, OPTIMISATION LOGICIELLE, REFRROIDISSEMENT, ETC., TOUTES LES PISTES SONT EXPLORÉES POUR RÉDUIRE LEUR VORACITÉ.

LE PROCHAIN DÉFI : LA MAÎTRISE DE L'ÉNERGIE



FUJITSU

Le calcul haute performance est un sport de haut niveau. Pour battre de nouveaux records, les « athlètes » de cette discipline ont aujourd'hui une priorité : réduire par tous les moyens leur alimentation... en électricité. La voracité énergétique des machines est en effet devenue le frein majeur au développement de supercalculateurs plus puissants.

À l'heure actuelle, les supercalculateurs ont besoin d'une très grosse quantité d'électricité. Champion du monde de calcul intensif depuis juin dernier, le système japonais K, co-développé par le

fabricant Fujitsu et l'Institut avancé de sciences informatiques de Kobe, au Japon, consomme quasiment 10 mégawatts à plein régime, soit la consommation électrique annuelle hors chauffage d'environ 5 000 foyers !

Grâce à cette énergie, il a affiché un score de 8,16 millions de milliards d'opérations par seconde, soit 8,16 pétaflops⁽¹⁾ au test Linpack⁽²⁾, prenant ainsi la première place du Top 500, un classement international des supercalculateurs qui, bien que contesté, reste très utilisé. En valeur absolue, la consommation du super K est impressionnante et, parado-

xalement, il affiche aussi l'une des meilleures efficacités énergétiques de sa catégorie. Sa « performance par watt » est de 824 mégaflops⁽¹⁾ par watt (8,16 pétaflops rapportés à 9,9 mégawatts), alors que l'efficacité énergétique moyenne des dix premiers supercalculateurs du Top 500 est de seulement 463 mégaflops par watt.

Ce bon résultat est toutefois encore insuffisant pour envisager la construction d'une machine capable d'atteindre le prochain objectif que se sont fixés en 2009 les acteurs du domaine : l'exaflop⁽¹⁾, c'est-à-dire le milliard de milliards d'opérations par seconde.

Équipé de 68 544 processeurs, l'ordinateur K est le supercalculateur le plus puissant du monde depuis juin dernier. Son efficacité énergétique est aussi l'une des meilleures du monde.



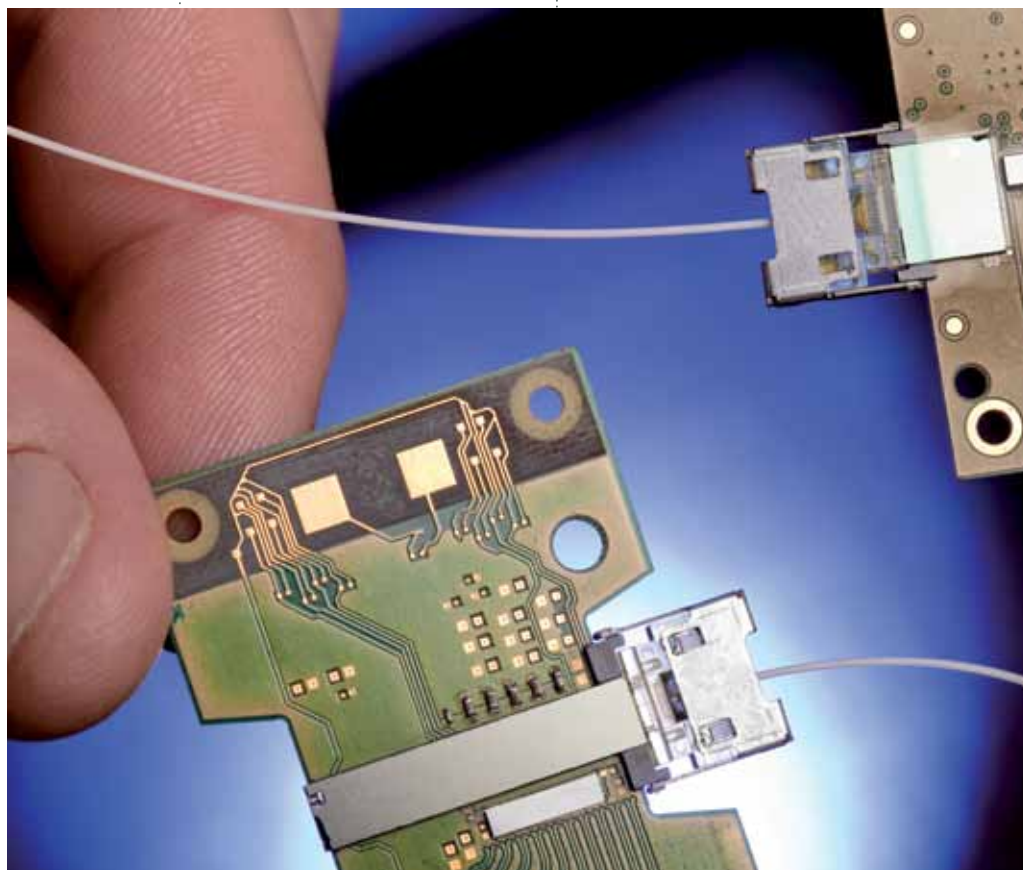
1. FLOPS
Le flops (Floating point Operations Per Second) est l'unité de mesure de la puissance des ordinateurs en nombre d'opérations par seconde. Un teraflops permet de faire mille milliards d'opérations par seconde (10^{12}) et un exaflops un milliard de milliards par seconde (10^{18}).

2. LINPACK
Ce test de performance mesure le temps nécessaire à un ordinateur pour résoudre un système d'équations linéaires de « n » équations à « n » inconnues.

« On estime que début 2012, les supercalculateurs les plus puissants délivreront 10 pétaflops en consommant 10 mégawatts, soit environ un pétaflop par mégawatt, explique Franck Cappello, codirecteur du laboratoire commun à l'Inria et à l'université de l'Illinois (États-Unis), spécialiste de la course à l'exaflop. En extrapolant ces résultats, un supercalculateur exaflopique consommerait 1 000 mégawatts. Ce n'est pas acceptable. »

Même si la demande pour des calculateurs capables d'atteindre l'exaflop est très forte, le coût d'un tel niveau de consommation énergétique n'est pas viable économiquement. Pour un supercalculateur qui peut « brûler » jusqu'à 1 000 mégawatts à plein régime – soit autant qu'une navette spatiale au décollage –, la facture d'électricité annuelle pourrait dépasser les 500 millions d'euros en France, soit environ deux fois le coût du supercalculateur lui-même : entre 200 et 300 millions d'euros. « L'objectif est de concevoir d'ici à 2018 un supercalculateur exaflopique consommant seulement 20 mégawatts car cela correspond à la capacité maximale des infrastructures ayant besoin d'accueillir une telle machine, explique Franck Capello. Certains pensent toutefois qu'il sera difficile de rester en dessous des 50 mégawatts. »

Avec un objectif de 20 mégawatts, l'efficacité énergétique d'un supercalculateur exaflopique serait de 50 000 mégaflops par watt ! Pourra-t-on multiplier la puissance des supercalculateurs actuels par 100 tout en améliorant leur efficacité énergétique d'un facteur 50 ? Pour relever ce défi, les chercheurs font feu de tout bois. Un premier axe d'amélioration est la mise au point de microprocesseurs et de mémoires plus économes en énergie. Pour réaliser des milliards d'opé-



Conçu par Intel, ce prototype de circuit en silicium produit un laser qui permet d'échanger jusqu'à 50 gigabits de données par seconde ; une aubaine pour les fabricants de supercalculateurs.

rations par seconde, les supercalculateurs ont en effet besoin de toujours plus de processeurs et de mémoires qui contiennent d'innombrables transistors ayant chacun besoin de courant électrique pour s'ouvrir ou se fermer de concert à chaque cycle d'opérations. Il leur faut une quantité d'énergie phénoménale – une grande partie se dissipe qui plus est sous forme de chaleur par effet Joule –, ce qui nécessite l'installation de systèmes de refroidissement eux aussi gourmands en énergie.

Plus de 500 000 cœurs

Depuis trente ans, les industriels qui fabriquent ces composants

« Si on divise par deux le voltage, il faut multiplier par quatre le nombre de cœurs pour conserver le même niveau de performance »

(Intel, IBM, AMD, Fujitsu, Nvidia, etc.) augmentent la finesse de gravure, ou le diamètre du plus petit fil reliant deux composants d'un circuit. Ils ont aussi mis au point des alliages qui ont permis d'augmenter les fréquences de commutation des transistors tout en limi-

L'EFFICACITÉ ÉNERGÉTIQUE DES DIX PREMIERS SUPERCALCULATEURS

RG	NOM	FABRICANT	SITE	PUISANCE DE CALCUL (PÉTAFLIPS)	CONSOMMATION ÉLECTRIQUE (MÉGAWATTS)	EFFICACITÉ ÉNERGÉTIQUE (MÉGAFLOPS / W)
1	Ordinateur K	Fujitsu	Japon	8,16	9,9	824,6
2	Tianhe-1A	NUDT	Chine	2,57	4,04	635,1
3	Jaguar	Cray	États-Unis	1,76	6,95	253,1
4	Nebulae	Dawning	Chine	1,27	2,58	492,6
5	TSUBAME 2.0	NEC/HP	Japon	1,19	1,4	852,3
6	Cielo	Cray	États-Unis	1,11	3,98	278,9
7	Pleiades	SGI	États-Unis	1,09	4,10	265,2
8	Hopper	Cray	États-Unis	1,05	2,91	362,2
9	Tera 100	Bull	France	1,05	4,59	228,8
10	Roadrunner	IBM	États-Unis	1,04	2,35	444,3



2010 JEFFREY TSENG @ INTOUCHSTUDIOS.COM

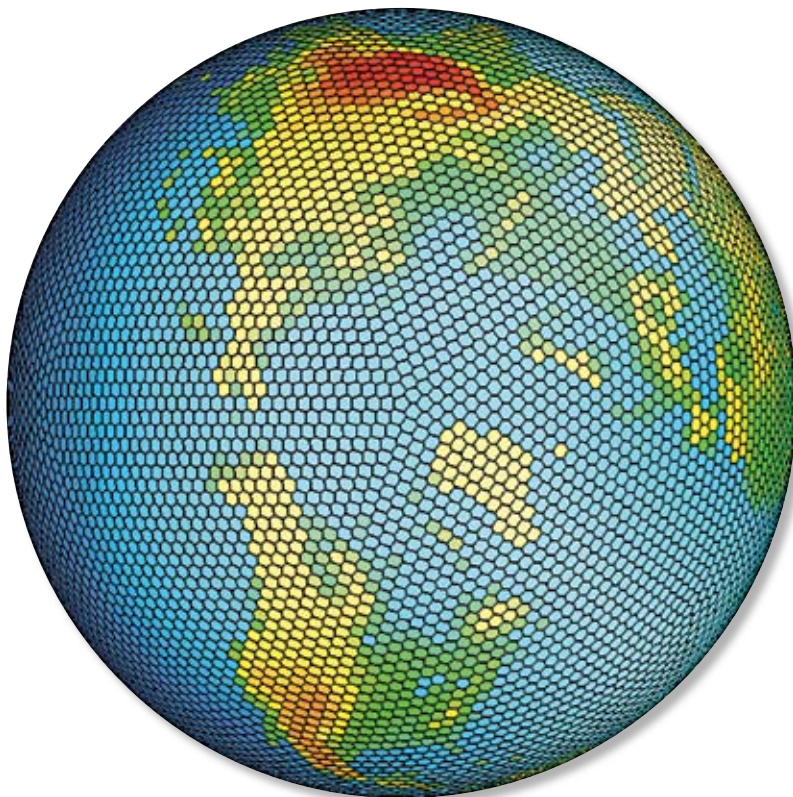
tant les voltages des puces. Mais la chaleur dégagée sur des surfaces toujours plus petites a atteint des niveaux ingérables dès 2004, et les fondeurs ont dû se résoudre à plafonner les fréquences de leurs puces qui dépassent désormais rarement les 4 GHz. Dans un supercalculateur, elle oscille même généralement entre 1 et 3 GHz.

La vitesse ne pouvant plus augmenter, les progrès en miniaturisation ont été mis au service du parallélisme. Chaque microprocesseur contient désormais plusieurs cœurs qui sont des unités de traitement capables de travailler de manière autonome et en parallèle. L'ordinateur K utilise par exemple 68 544 processeurs Sparc64 VIIIx à 8 cœurs, gravés en 45 nm et cadencés à 2 GHz, ce qui porte à 548 352 le nombre total de ses cœurs. Et ce n'est qu'un début. « On prévoit qu'en 2018, les puces seront gravées en 8 nm et que les processeurs contiendront plus de 1 000 cœurs », anticipe Franck Cappello.

Cette prolifération des cœurs permet notamment d'améliorer l'efficacité énergétique en diminuant le voltage des puces. Une technique baptisée « *voltage scaling* » en anglais. En abaissant le voltage d'un processeur, on réduit sa consommation, mais on abaisse aussi son niveau de performance et il faut plus de cœurs pour compenser. Si on divise par deux le

LE CO-DESIGN À LA RESCOURSSE

Vu les rendements énergétiques des supercalculateurs généralistes actuels, certains chercheurs pensent qu'il faudra forcément envisager du co-design pour mettre au point une machine exaflopique consommant moins de 50 mégawatts. Le co-design consiste à concevoir les architectures matérielles des machines en fonction de l'application qu'elles doivent faire tourner et donc à créer des supercalculateurs spécialisés cherchant à offrir les meilleures performances et/ou la meilleure efficacité énergétique. Il est très probable que le premier supercalculateur exaflopique sera un supercalculateur spécialisé. L'inconvénient de ces machines est bien sûr leur trop grande spécialisation.



Le projet Green Flash vise à concevoir un simulateur climatique capable d'atteindre 200 pétaflops en consommant 4 mégawatts, soit une efficacité énergétique de 50 000 mégaflops par watt. Il mise pour cela sur une architecture massivement parallèle, rassemblant 20 millions de cœurs, et réalisée sur mesure par rapport à l'application cible. Un prototype a montré la pertinence de cette approche à la fin 2010, mais il reste à trouver des financements, estimés à 75 millions de dollars, pour fabriquer l'ensemble.

voltage, il faut multiplier par quatre le nombre de cœurs pour conserver le même niveau de performance. Il s'agit ici d'un axe majeur pour réduire la facture d'électricité. « En 2018, les supercalculateurs compteront plusieurs centaines de millions de cœurs, contre 300 000 aujourd'hui », annonce Franck Cappello.

Toujours trop de consommation

Autre levier pour gagner en efficacité énergétique : l'amélioration des mémoires et des liaisons de communication entre composants. Actuellement, le déplacement des données entre les processeurs et les mémoires consomme en effet plus d'énergie que les traitements réalisés au sein des

processeurs ! Là aussi, les concepteurs misent sur les progrès des fabricants de composants.

Une première innovation espiérée par les fabricants de supercalculateurs est ici la mise en œuvre de liaisons optiques, en lieu et place des câbles en cuivre, pour relier armoires et cartes électroniques, ainsi qu'au sein des puces en remplacement de certains traces sur les circuits imprimés.

En utilisant des photons plutôt que des électrons pour transporter l'information, les liaisons photoniques créées par des puces en silicium émettrices/réceptrices d'un laser permettent d'atteindre en laboratoire des débits très élevés (plus de 50 Gbits/sec) et promettent de réduire la consomma-



3. MÉMOIRE À CHANGEMENT DE PHASE

Les mémoires à changement de phase (ou PRAM, pour Phase-Change Random Access Memory) enregistrent l'information dans un matériau vitreux qui change d'état quand un courant électrique est appliqué et elles n'ont pas besoin d'être alimentées en permanence.

4. MEMRISTOR

Le memristor (ou memristance) est un composant électronique passif dont la valeur de la résistance change de façon permanente sous l'effet d'une impulsion électrique. Il est utilisé pour la conception de mémoires dites RRAM ou ReRam pour Resistive Random Access Memory qui n'ont pas besoin d'être alimentées en permanence.

tion énergétique d'un facteur 10. «La mise en œuvre d'un réseau de communication optique de bout en bout sera un saut technologique majeur dans la course à l'exaflop», souligne Patrick Demichel, architecte système spécialiste du calcul intensif chez HP. Les fabricants développent aussi des puces mémoire permettant un empilement en trois dimensions grâce à des interfaces de communication verticales. Concrètement, il sera possible de les empiler au-dessus des processeurs et donc de réduire les distances entre composants.

Mais c'est surtout l'avènement des mémoires non volatiles, comme les mémoires à changement de phase⁽³⁾ ou les memristors⁽⁴⁾, qui permettrait de réduire considérablement la facture d'électricité. Contrairement à la mémoire DRAM actuelle, les mémoires non volatiles n'ont pas besoin d'être alimentées en permanence. «Si elles sont suffisamment performantes, elles permettront de revoir l'architecture des machines et de simplifier les dispositifs de tolérance de

panne qui auront moins de sauvegardes à effectuer sur disque dur ou SSD⁽⁵⁾, et pourront se contenter de micropoints de contrôle», explique Franck Cappello.

Si l'on se contente de l'évolution technologique des composants sans remise en cause des architectures, les experts estiment que la consommation hors système de refroidissement d'un supercalculateur exaflopique pourrait atteindre 150 mégawatts en 2018. Les processeurs consommeraient 50 mégawatts, la mémoire 80 mégawatts et le réseau 20 mégawatts. Il faudrait ensuite ajouter la consommation du système de refroidissement soit 50 à 70 % d'énergie supplémentaire pour un total dépassant les 200 mégawatts. C'est encore beaucoup trop.

Un accélérateur

Pour aller plus loin, l'architecture du supercalculateur sera un choix déterminant. Les concepteurs réfléchissent à la nature des unités de traitement à utiliser au sein des supercalculateurs.

Deux tendances s'affrontent. Certains privilégient des machines utilisant un seul modèle de processeur généraliste – à l'instar de l'ordinateur K –, tandis que d'autres combinent des processeurs généralistes (des CPU pour Central Processing Unit) et des processeurs graphiques (GPU⁽⁶⁾) dans des machines «hybrides» – ainsi la deuxième machine la plus puissante du monde, le Tianhe-1A, installée en Chine à Tianjin, qui atteint 2,5 pétaflops pour 4 mégawatts.

Utilisés ici pour effectuer des calculs et non pas pour afficher des graphismes, les GPU jouent le rôle d'accélérateur dans certaines applications et permettent d'améliorer l'efficacité énergétique globale de la machine. Outre leur spécialisation, leur point faible à l'heure actuelle est qu'ils ne communiquent pas rapidement avec les autres GPU car ils ont besoin des CPU qui jouent les intermédiaires. Pour les applications où les processeurs communiquent beaucoup entre eux, cette approche n'est donc pas pertinente. La si-

LE REFROIDISSEMENT, OBJET DE TOUTES LES ATTENTIONS

Le système de refroidissement qui permet de garantir le bon fonctionnement d'un supercalculateur augmente en général la facture d'électricité de 50 à 75 %. Les concepteurs multiplient donc les innovations pour concevoir des systèmes de climatisation plus efficaces. La priorité des fabricants est de fonctionner en free-cooling, c'est-à-dire en faisant circuler l'air ambiant sans aide de pompe à chaleur. Quand il est possible, le choix d'un site situé dans une région froide est à privilégier mais le free-cooling reste un objectif impossible à atteindre en permanence pour un supercalculateur exaflopique. Les fabricants tentent donc de trouver des solutions complémentaires les moins gourmandes possibles. Les portes des armoires qui contiennent les grappes de processeurs contiennent par exemple des circuits d'eau glacée. Les cartes électroniques de certains fabricants sont parcourues par des circuits de refroidissement et IBM prépare même des puces parcourues et refroidies par des microcanaux de fluide réfrigérant. Malgré tous ces efforts, certains experts pensent que les propriétaires de supercalculateurs devront revendre la chaleur récupérée pour équilibrer leur modèle économique, une idée séduisante mais pas forcément évidente à concrétiser dans la pratique.



Les portes des armoires du plus puissant supercalculateur français, le Tera 100 (9^e rang mondial), de Bull, contiennent un échangeur, des ventilateurs et une tuyauterie d'eau glacée reliée au faux plancher totalisant 1 km. Chaque porte dissipe 40 kW par armoire.

CEA



tuation pourrait venir à changer. «À termes, on ne distinguera vraisemblablement plus les CPU et les GPU car les processeurs contiendront à la fois des cœurs généralistes et des cœurs spécialisés qui joueront le rôle d'accélérateurs», anticipe Franck Cappello.

« Si on divise par deux le voltage, il faut multiplier par quatre le nombre de cœurs pour conserver le même niveau de performance »

Vu la complexité des futurs supercalculateurs, le logiciel tiendra demain une place primordiale. «Dans la course à l'exaflop, le principal problème scientifique sera de trouver de bonnes méthodes de programmation et de nouveaux modèles mathématiques, prédit Serge Petiton, responsable de l'équipe MAP (Méthodologie et Algorithmique Parallèle pour le calcul scientifique)



5. SSD

Un disque SSD (pour Solid-state drive) est un matériel de stockage de données constitué de mémoires Flash.

6. GPU

Un processeur graphique, ou Graphics Processing Unit en anglais, est un processeur dédié au calcul de l'affichage. Il est capable d'effectuer plusieurs tâches en parallèle.

au Laboratoire d'informatique fondamentale de Lille. Je pense que nous nous trouverons face à un mur au-delà des 100 pétaflops et qu'il faudra absolument inventer de nouveaux paradigmes.»

Aujourd'hui déjà, il est de plus en plus difficile de paralléliser les calculs, c'est-à-dire de les diviser en sous-calculs exécutés par les cœurs des processeurs. Le logiciel doit en outre privilégier la localité, donc limiter les mouvements de données entre processeurs. Cela se traduit par l'apparition de nouvelles disciplines algorithmiques comme la «Communication Avoiding Algorithm», dont l'objectif est de permettre au processeur de travailler au maximum en autarcie.

Comme il n'est jamais possible d'avoir en permanence un parallélisme parfait, le logiciel doit désormais être capable d'éteindre les ressources matérielles non utilisées. Ce rôle est déjà dévolu aux systèmes d'exploitation et à certains matériels comme le processeur qui peut ralentir ou désactiver certains de ses cœurs. Aussi les applications

sont conçues d'emblée pour économiser l'énergie. «Le programmeur a maintenant trois critères à gérer : le nombre d'itérations, la durée de chaque itération et l'énergie globale associée à l'exécution de toutes ces itérations, explique Serge Petiton. Et selon les besoins, ses algorithmes s'emploient à réduire le temps de calcul ou la facture énergétique.»

Avec des centaines de milliers de processeurs, il est de plus en plus difficile de prévoir quelle sera la meilleure méthode de calcul pour réduire la facture électrique. Selon Serge Petiton : «Ce n'est pas un problème déterministe. C'est pourquoi nous mettons en place des techniques d'autotuning qui consistent à faire évoluer automatiquement en temps réel les paramètres de calcul afin de réduire la consommation énergétique.» En optimisant les systèmes d'exploitation, les langages, les compilateurs et les applications, on pourrait ainsi économiser 10 à 20 % d'énergie.

Pour Monsieur tout-le-monde

Les gains obtenus pourront être importants mais la complexité des architectures nécessitera de renforcer les dispositifs de tolérance aux pannes (systèmes de corrections d'erreur, sauvegardes des registres...). «Nous estimons que cela induira un surcoût énergétique et, au final, les mécanismes de tolérance de pannes devraient drainer un tiers de l'énergie dépensée par le supercalculateur contre 20 % à ce jour», imagine Franck Cappello.

Pour trouver des solutions à tous ces problèmes, l'ingéniosité des chercheurs sera indispensable. Ainsi, Serge Petiton estime que «pour concevoir un calculateur exaflopique, il faudra des équipes pluridisciplinaires de plusieurs centaines de personnes, formées à la problématique de l'énergie; ce sera un vrai challenge en matière de recrutement et de formation.»

Mais le jeu en vaut la chandelle. En mettant au point des machines atteignant l'exaflop, une multitude de problèmes ayant un fort impact sur la société pourront être résolus, comme les prévisions climatiques ou le criblage des médicaments. Et de la même manière que les innovations en Formule 1 se retrouvent sur les voitures de Monsieur tout-le-monde, les progrès réalisés dans le domaine des supercalculateurs devraient permettre de réduire considérablement la consommation électrique des ordinateurs et des équipements high-tech grand public. Vivement les prochains grands challenges! ●

CONSTANTIN LENSELLE



TOUT EN CONSERVANT L'UTILISATION DE COMPOSANTS STANDARD, HP FERA APPEL À L'OPTIQUE POUR RÉDUIRE CONSIDÉRABLEMENT LA CONSOMMATION D'ÉNERGIE DES MACHINES ET PARVENIR À L'EXASCALE.

JEAN-LUC ASSOR : « HP A TRACÉ LA VOIE JUSQU'EN 2020 »



Jean-Luc Assor est responsable du marché Engineering (aéronautique, automobile et semi-conducteurs) au sein de la Division HPC chez HP Corp.

Dans le calcul haute performance, HP se distingue en proposant des solutions dotées de composants standard : n'est-ce pas paradoxal ?

Nous ne voulons pas rentrer dans une logique de performance à tout prix, à l'image d'une Formule 1 conçue uniquement pour battre des records. Notre ambition est de partir de solutions conformes aux standards de l'industrie informatique pour les rendre plus performantes. C'est ainsi que HP se distingue parmi la dizaine d'acteurs présents sur le calcul haute performance. C'est une stratégie payante d'un point de vue commercial car nous élargissons le marché des

supercalculateurs. Ainsi, notre premier débouché n'est pas la recherche, mais les services : banques, assurances, effets spéciaux pour imagerie numérique... En jouant sur le grand volume de composants standard, nous parvenons à faire baisser les prix. Cette approche est par ailleurs une garantie pour nos clients de notre engagement sur le long terme. Adopter des composants standard, c'est un gage de pérennité et de fiabilité.

Le risque n'est-il pas de décrocher de la course à la puissance ?

Non, car nous travaillons sur des solutions innovantes pour les prochaines générations. Seulement, nous nous assurons que les technologies les plus performantes deviennent des standards de l'industrie avant de les déployer. Prenons l'exemple des cartes graphiques (« many cores ») : elles permettent de démultiplier les performances des serveurs, mais à la condition que l'on sache les programmer de façon pérenne et que l'on soit sûr d'arriver à des calculs numériquement justes. Or, ces deux conditions ne sont remplies que depuis 2010. Maintenant que la technologie est fiable, nous la mettons en œuvre. HP est d'ailleurs le premier revendeur mondial de solutions « many cores » avec une gamme très large, qui va des stations de travail aux serveurs haute densité.

Les architectures actuelles seront-elles adaptables pour parvenir à l'exascale ?

Non : pour la première fois depuis dix ans, nous devons redéfinir les principes de

l'informatique et l'architecture des ordinateurs. Les technologies à base de cuivre atteignent leurs limites pour des questions énergétiques : afin de faire fonctionner une machine exascale aujourd'hui, il faudrait l'alimenter par une ou deux centrales nucléaires ! La seule solution est de repenser une bonne partie de l'architecture des ordinateurs pour gagner d'ici à 2020 un facteur 5 ou 10 sur la consommation énergétique.

Investissez-vous une nouvelle voie pour y parvenir ?

Oui, il faudra en effet faire appel à l'optique, et à l'échelle du nanomètre, que ce soit pour la communication entre les processeurs comme pour les échanges avec la mémoire ; car augmenter la finesse de gravure des logiques de calcul ne suffira pas. Nous avons développé le concept de « ring resonator » ou résonateur en réseau : une structure de quelques micromètres qui permet de moduler la lumière et de passer du domaine électrique au domaine optique en consommant très peu d'énergie. Enfin, pour avoir un système résistant aux pannes, nous développons le « memristor », qui offrira une grande quantité de mémoire non volatile et pour un faible coût, tout en ayant de bonnes caractéristiques. Il pourra aussi être utilisée en mémoire de stockage. Avec les brevets déposés par HP, la route est tracée au moins jusqu'en 2020.



