



ÉDITORIAL
Denis Girou

L'année 2015 a été marquée par l'extension des capacités disques des configurations de calcul

Après l'ajout, fin 2014, de 32 000 cœurs supplémentaires au calculateur Blue Gene/Q, l'année 2015 a été marquée par l'extension des capacités disques des configurations de calcul, avec l'ajout de 3 Po supplémentaires aux 2 Po disponibles depuis 2012, via des disques de nouvelle génération aux performances nettement augmentées.

Après le remplacement en 2014 de la machine d'archives qui pilote notre robotique de cartouches magnétiques, c'est donc d'un environnement largement étendu dont nos utilisateurs vont pouvoir bénéficier jusqu'à la fin de vie de la génération actuelle des supercalculateurs Turing (Blue Gene/Q) et Ada (x3750), qui ont passé leur troisième année d'exploitation. Durant celle-ci, ils ont encore fait preuve d'une fiabilité remarquable, avec à nouveau pour chacun une disponibilité de l'ordre de 99,5 %, hors les arrêts programmés pour des mises à niveau des systèmes. En page 11, nous donnons les bilans des consommations d'heures par disciplines scientifiques en 2015, ainsi que les attributions de la première session de 2016.

Dans les pages 2 à 9, figurent les comptes rendus de quatre projets qui ont pu disposer en 2015 d'allocations spécifiques sur la machine Turing, dans le cadre de l'extension de cette configuration. On peut y voir, dans des disciplines différentes allant de la physique des plasmas à la physico-chimie,

de l'étude des écoulements turbulents à celle de protéines en cause dans certaines maladies neurodégénératives, l'apport de l'accès à des configurations massivement parallèles, dont l'usage simultané d'un très grand nombre de cœurs de calcul permet des simulations inenvisageables il y a encore quelques années.

Toutefois, puisque notre configuration actuelle de supercalculateurs a dépassé sa mi-vie, il importe d'ores et déjà de préparer l'avenir et d'envisager, à l'horizon de leur remplacement, les choix technologiques qui seront arrivés à maturité en 2018. C'est pour préparer cette transition, tout comme celles qui auront lieu à leur tour dans les autres centres nationaux, et à un moment où l'on sait que des évolutions technologiques fortes sont nécessaires pour continuer à garantir la croissance des puissances disponibles et permettre l'arrivée d'ici quelques années des premières machines exaflopiques, que s'est mise en place, sous la coordination de GENCI, une cellule nationale de veille technologique. L'une de ses actions a été l'acquisition de deux plates-formes, prototypes des architectures à venir, l'une chez nos collègues du CINES et l'autre à l'IDRIS (voir en page 10).

Du côté des projets européens, en attendant le démarrage de la seconde phase de PRACE, dite PRACE 2, les projets d'implémentation continuent, avec 4IP qui en est à mi-parcours et qui sera suivi par 5IP entre mai 2017 et avril 2019. Mais dans l'organisation globale repensée par la commission européenne dans le cadre du programme H2020, ce sont à présent trois piliers qui soutiennent et propulsent la construction progressive de l'Europe du calcul de haute performance : à côté de PRACE pour les infrastructures, se placent aujourd'hui ETP4HPC pour les technologies et les centres d'excellence pour les disciplines scientifiques. Suite à un appel à projets en 2014, huit centres d'excellence, eux-mêmes des partenariats entre plusieurs organismes, se sont constitués depuis octobre 2015. Ainsi qu'il est évoqué en page 10, l'IDRIS est associé à deux d'entre eux, dans le domaine de l'énergie et dans celui de la physico-chimie et de la physique des matériaux.

sommaire

- 1 • Éditorial
- 2 et 3 • Propagation d'ondes en milieu hétérogène et écoulements en compresseur multi-étages
- 4 et 5 • SMILEI - Amplification Brillouin de faisceaux laser ultra-courts

- 6 et 7 • Effets hydrodynamiques sur l'agrégation de l'amyloïde Aβ
- 8 et 9 • Solubilité du CO₂ dans les carbonates fondus
- 10 • Centres d'excellence européens
- 10 • Installation d'un prototype OpenPOWER à l'IDRIS
- 10 • Conférence Supercomputing 2015
- 11 • Bilan des consommations 2015 et allocations 2016
- 12 • Bilan des formations IDRIS 2015
- 12 • Calendrier des formations

Calcul haute performance pour la propagation d'ondes en milieu hétérogène et pour l'écoulement en compresseur multi-étages

1 TOTAL E&P Research and Technology USA, 2 INPT-IRIT, University of Toulouse and ENSEEIHT, 3 CERFACS

L'évolution du calcul haute performance permet d'aborder des problématiques de plus en plus complexes et réalistes à condition d'adapter, d'optimiser et de valider les codes sur les nouvelles architectures disponibles. Dans ce grand challenge, le CERFACS a proposé d'effectuer une double étude : deux solveurs ont été optimisés et validés, le premier permettant l'analyse de la propagation d'ondes en milieu hétérogène, le second étant adapté à la simulation d'écoulements fluides. En première partie, nous présenterons l'analyse du comportement de la méthode de résolution itérative pour la propagation d'ondes proposée dans [1] en fonction de la fréquence sur un très grand nombre de cœurs de calcul. En seconde partie, nous détaillerons les travaux effectués cette fois-ci sur la simulation d'écoulements extrêmement complexes dans un compresseur aéronautique.

Analyse de performance pour la simulation de la propagation d'ondes en milieu hétérogène en sismique

(H. Calandra¹, S. Gratton², X. Pinel³ et X. Vasseur³)

La simulation de phénomènes de propagation d'ondes en milieu hétérogène en sismique requiert généralement la résolution d'une équation aux dérivées partielles dite de Helmholtz. Cette activité concerne de multiples champs applicatifs et, de par sa complexité, représente un domaine très actif de recherche au niveau algorithmique, puisque des systèmes linéaires comportant des matrices indéfinies de très grande taille doivent être résolus dans un environnement massivement parallèle. Ce challenge nous a permis d'analyser sur Turing le comportement de la méthode de résolution itérative proposée dans [1] en fonction de la fréquence f .

Le tableau 1 synthétise les résultats obtenus relatifs à un cas test partagé par la communauté en sismique nommé EAGE/SEG Overthrust. Le cas $f=32$ Hz conduit à un système

f (Hz)	Maillage	Threads	Prec	T (s)	M (To)
4	512 x 512 x 128	131072	6	3	0,03
8	1071 x 1071 x 256	131072	13	10	0,16
12	1792 x 1792 x 512	131072	19	30	0,72
16	2560 x 2560 x 640	131072	25	77	1,74
20	3584 x 3584 x 896	131072	35	269	4,5
24	4352 x 4352 x 1024	131072	41	507	7,5
28	5376 x 5376 x 1280	131072	50	1155	13,9
32	6144 x 6144 x 1536	131072	62	2217	21,5

Tableau 1 - Analyse de complexité (EAGE/SEG Overthrust) en fonction de la fréquence. Méthode de Krylov préconditionnée par une méthode multigrille géométrique pour la résolution de l'équation d'Helmholtz en milieu hétérogène. Prec dénote le nombre d'applications du préconditionnement, T le temps de calcul total en secondes et M la mémoire totale requise en To.

linéaire comportant approximativement 58 milliards d'inconnues (l'arithmétique complexe simple précision a été utilisée), calcul que nous pouvons véritablement considérer comme frontière. Nous remarquons l'efficacité de l'approche développée conduisant à un nombre réduit d'applications du préconditionnement, y compris à haute fréquence. Si N dénombre le nombre total d'inconnues, le temps de calcul T se comporte comme $N^{1,28}$. De plus, nous notons que la mémoire totale croît linéairement avec le nombre d'inconnues car aucune factorisation directe ou approchée n'est requise au sein de la méthode numérique. Ce challenge effectué sur le calculateur GENCI Turing de l'IDRIS a été l'occasion de réaliser des calculs frontières utilisant plus de 130 000 *threads* de calcul pour la résolution de l'équation d'Helmholtz et de démontrer l'efficacité de la méthode proposée en terme de complexité vis-à-vis de la fréquence. Dans un futur proche, la réflexion se portera sur les adaptations algorithmiques à réaliser pour tirer pleinement parti des futures machines exascale.

Simulation aux grandes échelles et validation pour les écoulements dans les compresseurs aéronautiques

(J. de Laborderie³, F. Duchaine³, L. Gicquel³, O. Vermorel³, G. Staffelbach³)

Les compresseurs présents dans les moteurs aéronautiques des avions et des hélicoptères représentent un composant critique en terme de rendement et de stabilité. Afin de permettre une avancée technologique sur la conception des futurs compresseurs, il est nécessaire de pouvoir réaliser des simulations aérodynamiques haute-fidélité de ces machines. L'écoulement se développant dans un compresseur est très complexe puisqu'il est instationnaire, possède une large gamme de structures turbulentes et dépend fortement des effets technologiques. Le code de simulation aérodynamique doit donc être capable de prendre en compte toutes ces caractéristiques.

L'objectif de l'étude consiste à démontrer la faisabilité d'une Simulation aux Grandes Échelles (SGE) d'un compresseur axial multi-étages avec le code TurboAVBP [2]. TurboAVBP est basé sur le couplage de plusieurs instances du même code AVBP [3], un code SGE résolvant les écoulements compressibles et réactifs. L'information est transférée entre ces instances, à chaque pas de temps, via un algorithme de couplage sur des maillages recouverts [2]. La configuration correspond au compresseur de recherche CREATE, conçu par SAFRAN-SNECMA et opéré au LMFA (École Centrale de Lyon).

Cette simulation est la première à être réalisée avec la technique de la SGE sur ce compresseur entier. Le compresseur haute pression CREATE possède 3,5 étages (7 roues). Son nombre d'aubes par roue étant multiple de 16, le domaine de calcul est limité à un secteur angulaire de $2\pi/16$. Ce domaine, illustré sur la figure 1 (a), comprend 37 aubages ainsi que les effets technologiques tels que les jeux de fonctionnement en tête de pale des rotors et la cavité de recirculation sous le premier stator.

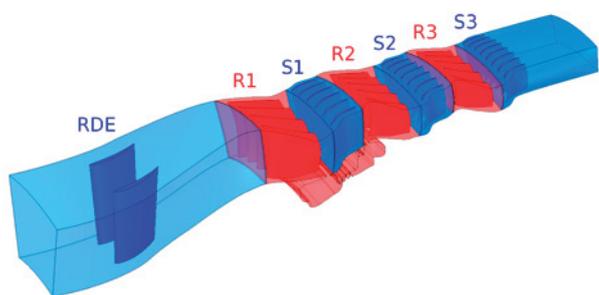


Figure 1 (a) - Domaine du calcul.

Le domaine est divisé en deux instances, l'une fixe (en bleu) comprenant la roue directrice d'entrée (RDE) et les stators S1, S2 et S3, et l'autre mobile (en rouge), constituée des rotors R1, R2 et R3 ainsi que de la cavité liée à R1 et R2. Trois maillages non-structurés ont été réalisés, qui contiennent respectivement 88 millions (M1), 194 millions (M2) et un milliard (M3) de cellules. L'une des difficultés de ces simulations consiste à obtenir un écoulement stabilisé, qui résulte d'un équilibre entre les conditions aérodynamiques imposées à l'entrée et à la sortie, le travail apporté par les rotors sur l'écoulement et les pertes calculées. Après une période transitoire, il a été observé une variation de débit inférieure à 0,1% en moyenne entre deux rotations successives, ce qui montre la convergence des simulations.

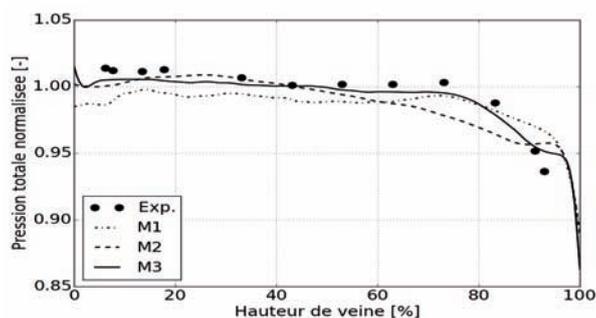


Figure 1 (b) - Évolution radiale de la pression totale moyennée derrière le 2^e rotor.

La prédiction des performances aérodynamiques globales au point de fonctionnement nominal est correcte, puisque l'écart avec le taux de pression mesuré en essais est inférieur à 2 % pour les 3 maillages. Les mesures détaillées effectuées à plusieurs endroits dans le compresseur permettent d'évaluer plus finement les simulations effectuées. La figure 1(b) présente l'évolution radiale de la pression totale, moyennée en temps et dans la direction circonférentielle derrière le deuxième rotor.

Sur ces graphes, le maillage le plus fin (M3) donne les meilleurs résultats par comparaison avec les données expérimentales. En particulier, dans la partie haute de la veine, seul M3 est capable de prédire correctement la forme et le niveau de la pression totale. Cela confirme la nécessité de discrétiser finement les jeux des rotors et la partie haute de la veine.

Afin d'illustrer les écoulements se développant dans la partie haute de la veine du compresseur, la figure 2 présente les structures turbulentes dans le dernier étage (R3-S3) au-dessus des pales du rotor. Les tourbillons créés par la fuite dans le jeu sont correctement capturés par la simulation haute-fidélité.

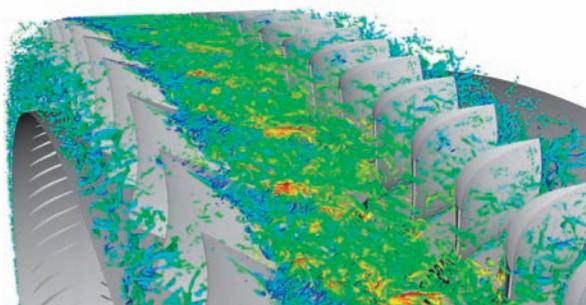


Figure 2 - Critère Q coloré par le nombre de Mach, dernier étage de CREATE, maillage M3.

Les ressources fournies par GENCI sur le calculateur Turing de l'IDRIS dans le cadre de ce grand challenge ont permis d'atteindre l'objectif de l'étude : des calculs SGE ont été effectués sur un compresseur multi-étages réaliste. De plus, les performances globales sont correctement prédites et les écoulements de jeux sont capturés. Cette étude se poursuivra, en cherchant en particulier à évaluer ces simulations avec des mesures instationnaires existantes.

[1] H. Calandra, S. Gratton, X. Pinel, X. Vasseur, An improved two-grid preconditioner for the solution of three-dimensional Helmholtz problems in heterogeneous media. *Numer. Linear Algebra Appl.*, 20, pp. 663-688 (2013).

[2] G. Wang, F. Duchaine, D. Papadogiannis, I. Duran, S. Moreau, L. Gicquel, 2014, An overset grid method for large eddy simulation of turbomachinery stages, *J. Comput. Phys.*, 274.

[3] <http://cerfacs.fr/avbp7x>

SMILEI - Amplification Brillouin de faisceaux laser ultra-courts

M. Chiaramello¹, J. Dérouillat², M. Flé³, C. Riconda¹, M. Grech¹

¹ Laboratoire d'Utilisation des Lasers Intenses, ² Maison de la Simulation

³ Institut du développement des ressources en informatique scientifique

Pour accompagner au mieux le formidable effort expérimental et technologique mené dans le cadre du futur laser multi-petawatt Apollon, plusieurs laboratoires du plateau de Saclay se sont fédérés autour du projet SMILEI visant à développer un nouvel outil de simulation dédié à l'interaction laser-plasma. SMILEI est un code *Particle-In-Cell* (PIC) développé dans un cadre *open source* et collaboratif et pensé pour prendre en compte au mieux les nouveaux paradigmes du calcul haute performance (HPC). SMILEI a ainsi été initié dans le cadre d'un co-développement entre physiciens des plasmas et spécialistes HPC. Le projet présenté ici a permis de réaliser les premières simulations massivement parallèles avec SMILEI en vue d'une étude physique originale visant à modéliser, dans des conditions réalistes, l'amplification de faisceaux laser ultra-courts via l'interaction laser-plasma.

Amplification de faisceaux laser ultra-courts dans un plasma

Un processus fondamental dans les plasmas est le couplage résonnant à trois ondes : deux ondes électromagnétiques sont couplées via une troisième onde (longitudinale) qui est un mode propre du plasma. Le plasma se comporte alors comme un milieu amplificateur transférant l'énergie d'une des ondes électromagnétiques vers l'autre onde électromagnétique. En particulier, si l'on fait croiser une impulsion longue, énergétique d'intensité pas trop élevée (onde dite de pompe) avec une impulsion courte de faible intensité (onde dite sonde), l'énergie de l'onde de pompe peut être en grande partie transférée à l'impulsion courte via l'excitation résonnante d'une onde plasma. Il est ainsi possible de créer des impulsions courtes de très grande intensité. L'avantage de ce type d'approche, en comparaison avec l'amplification conventionnelle dans les solides, est que le plasma peut supporter des ondes électromagnétiques de très grande intensité, et l'amplification plasma de faisceaux courts n'est donc pas contrainte par les problèmes d'endommagement optique. Cette approche est une voie prometteuse pour les optiques innovantes des futurs lasers de très grandes puissances (exawatt et au-delà).

Lorsque l'onde plasma est une onde plasma électronique, on parle d'amplification Raman. Lorsque l'onde plasma est une onde acoustique ionique, on parle d'amplification Brillouin. Pour les lasers de pompe à haute intensité, $I\lambda^2 > 10^{15} \text{ W/cm}^2\mu\text{m}^2$, le régime d'amplification Brillouin – dit de couplage fort – a été récemment mis en avant du fait que les propriétés de l'onde plasma sont gouvernées par les propriétés de la pompe, et permet d'aller vers des intensités de la sonde amplifiée particulièrement élevées. Aussi, l'onde excitée dans ce régime est un quasi-mode du plasma avec un spectre large de fréquences permettant un couplage amélioré entre pompe et sonde. Nous nous sommes ici intéressés à ce scénario, l'amplification Brillouin étant plus robuste que l'amplification Raman (plus sensible e.g. aux inhomogénéités du plasma) et ne requérant pas un décalage en fréquence de la sonde (expérimentalement contraignant).

Le projet proposé pour le Grand Défi

Le projet présenté ici a d'une part permis les premières simulations massivement parallèles avec SMILEI (Fig.1). Un très bon passage à l'échelle a été démontré sur CINES/Occigen (Bullx) jusqu'à 49 512 processeurs.

Sur Turing (Blue Gene/Q), une efficacité superlinéaire a été démontrée jusqu'à 262 144 *threads* OpenMP (65 536 tâches MPI sur 4 096 nœuds). Ce comportement s'explique par le traitement local des calculs relatifs aux particules, rendu possible par une méthode de tri mise en œuvre pour améliorer l'exploitation du cache et une parallélisation hybride MPI/OpenMP.

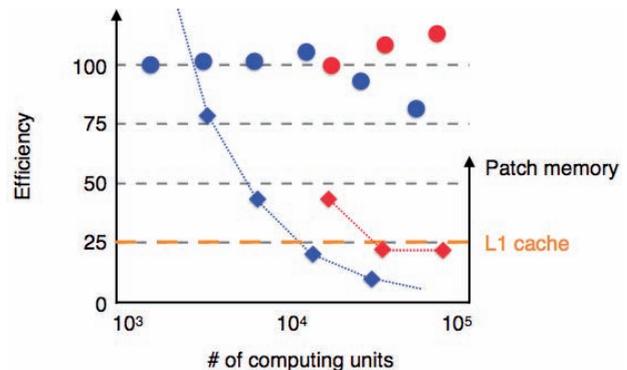


Figure 1 - Passage à l'échelle (strong scaling) du code SMILEI sur les machines CINES/Occigen (Bullx) en bleu, et IDRIS/Turing (Blue Gene/Q) en rouge.

D'autre part, ce projet a permis de réaliser les premières simulations PIC à 2 dimensions d'espace (2D) à grande échelle prenant en compte l'ensemble des conditions expérimentales actuelles. Simuler l'amplification Brillouin de faisceaux laser dans les conditions pertinentes pour les études expérimentales, i.e. à échelle réelle, est un défi qui requiert le recours aux machines massivement parallèles. En effet, on doit en même temps pouvoir décrire une large gamme d'échelles spatiales et temporelles. Le plasma est typiquement millimétrique, les tailles de faisceaux laser sont de quelques dizaines à une centaine de micromètres, et les longueurs en jeu dans l'interaction (longueur d'onde du laser et des ondes plasmas) sont micrométriques. Aussi, la physique évolue sur des temps de l'ordre de la dizaine de picosecondes (ps), les impulsions laser étant de l'ordre de quelques ps (quelques 100 fs) pour la pompe (pour la sonde) et les échelles électroniques et du cycle optique de quelques fs.

Afin de prendre en compte l'ensemble des phénomènes physiques mis en jeu (compétition entre instabilités Raman, Brillouin et filamentation, processus de saturation), une description cinétique est nécessaire et requiert donc l'utilisation de codes PIC. Enfin, seule une simulation à plusieurs dimensions d'espace (ici 2D) permet une description correcte de la physique mise en jeu. Bien que particulièrement exigeantes, de telles simulations sont aujourd'hui nécessaires du fait des récents efforts expérimentaux engagés – notamment au LULI – dans cette voie.

Focus faisceaux sur deux types de simulations d'amplification Brillouin

Dans les simulations présentées ici, la boîte de simulation a une taille de 1 mm x 0,5 mm pour une résolution spatiale de 33 nm dans les deux directions. Nous utilisons donc 30 720 x 15 360 cellules PIC. L'interaction laser-plasma est simulée sur une durée de 10 ps avec une résolution de $7,3 \cdot 10^{-2}$ fs, correspondant à 135 000 pas de temps. Le nombre total de macro-particules utilisées varie entre 16 et 23 milliards.

L'empreinte mémoire (vive) de chaque simulation est de ~ 1 To, et chaque simulation a généré ~ 1 To de données. Dans un premier temps, un ensemble de simulations a été réalisé dans la configuration typique des expériences actuelles mettant en jeu l'interaction d'une onde sonde avec une onde de pompe contra-propagative (Fig. 2). Des simulations préparatoires, à une dimension d'espace, avaient permis d'optimiser un certain nombre de paramètres comme le profil du plasma ou les profils temporels des impulsions laser. Les simulations en 2D, incluant de nouveaux effets comme l'autofocalisation ou la filamentation, nous ont permis d'optimiser un autre paramètre important : le retard entre les deux faisceaux lors de leur croisement dans le plasma. Ainsi, un retard de 2,5 ps apparaît comme optimal dans les conditions présentes, en très bon accord avec les expériences.

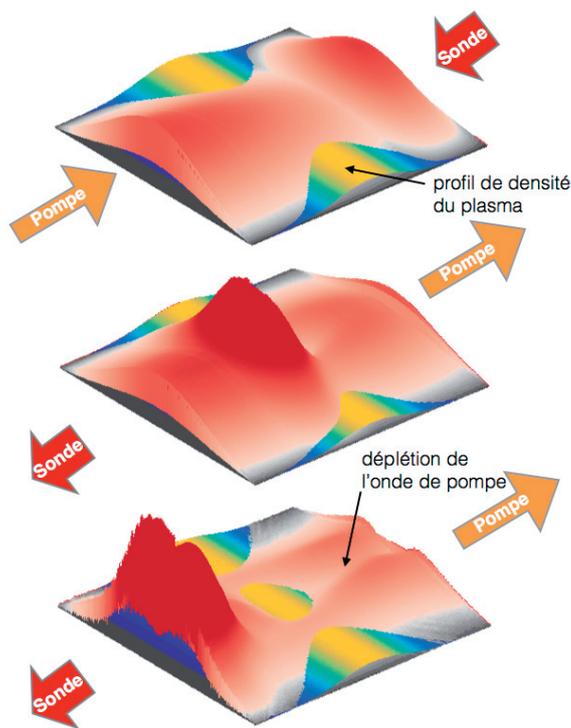


Figure 2 - Amplification Brillouin d'un pulse laser court (à droite) dans la configuration caractéristique des expériences avec une seule onde de pompe contra-propagative. La sonde entre par le bord droit de la simulation (image du haut à $t = 5,0$ ps), se propage et est amplifiée au sein du plasma (image centrale à $t = 7,6$ ps), et un faisceau court amplifié ($\times 4$) est obtenu en sortie de plasma (image du bas à $t = 9,6$ ps).

L'autre résultat important de ces simulations est la très bonne conservation du front d'onde de l'impulsion amplifiée. Il s'agit d'un résultat très important car il est envisagé de focaliser les impulsions courtes ainsi produites pour atteindre de plus grandes intensités, ce qui n'est possible que si le front d'onde n'est pas ou peu modifié pendant l'amplification. Dans les simulations antérieures, les tailles transverses avaient toujours été limitées par les contraintes numériques. Enfin, une simulation utilisant deux faisceaux de pompe a été réalisée. Ce nouveau schéma d'amplification a un très grand intérêt du point de vue expérimental car la configuration habituelle (pompe et onde contra-propagatives) est délicate à mettre en place du fait de possibles dommages dans les optiques. Un angle entre l'onde de pompe et la sonde peut être introduit, mais diminue

l'efficacité d'amplification. Dans le scénario proposé ici, deux ondes de pompe sont utilisées se propageant à un angle de 6° par rapport au faisceau sonde (Fig. 3), l'usage de deux faisceaux compensant en partie l'effet de l'angle. Cette simulation montre qu'une telle configuration permet une bonne amplification, et permet donc d'envisager des schémas alternatifs plus flexibles pour les expériences. Elle souligne aussi le rôle clef de la diffusion Brillouin dans l'amplification (Fig.4).

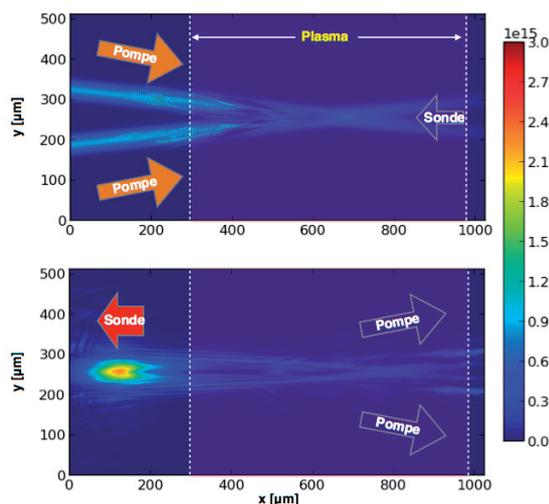


Figure 3 - Enveloppe des champs électromagnétiques extraits de la simulation avec deux faisceaux de pompe. En haut, les deux faisceaux de pompe ont commencé à interagir avec le plasma, alors que la sonde (à droite) n'est pas encore entrée dans la boîte de simulation. Dans la figure du bas, la sonde amplifiée ($\times 2,7$) a quitté le plasma.

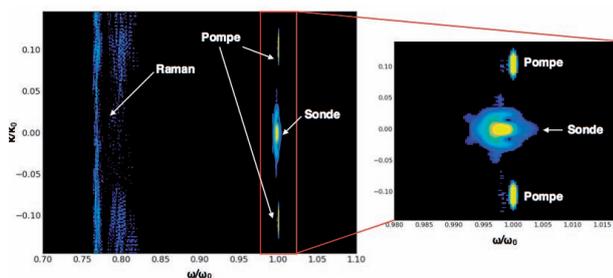


Figure 4 - L'analyse du spectre spatio-temporel des champs électromagnétiques mesurés au bord gauche (entrée de la pompe) de la boîte de simulation permet d'identifier les impulsions de pompe, l'onde de sonde et le rôle dominant de l'instabilité de diffusion Brillouin dans son amplification, ainsi que l'émission Raman spontanée.

Conclusions

Ce grand défi a permis, outre les premiers tests massivement parallèles du code SMILEI, la simulation réaliste d'expériences ayant un grand intérêt pour le développement de futures sources laser de très grande puissance. La robustesse du mécanisme d'amplification Brillouin a été démontrée en 2D sur de grandes échelles spatiales et temporelles.

Remerciements

Ce projet a été réalisé dans le cadre des Grands Défis GENCI 2015. Les auteurs remercient l'ensemble des partenaires du projet SMILEI et le soutien financier des labex PALM (PALM-ANR-10-LABX-39) et PLAS@PAR (ANR-11-IDEX-0004-02).

Effets hydrodynamiques sur l'agrégation de l'amyloïde A β

Mara Chiricotto¹, Simone Melchionna², Philippe Derreumaux¹ et Fabio Sterpone^{1*}

1. CNRS (UPR9080) – Institut de Biologie Physico-Chimique. Université de Paris Sorbonne Cité et Paris Sciences et Lettres, Laboratoire de Biochimie Théorique, Paris, France. 2. CNR – Institute of Complex System, Rome, Italie

L'agrégation de protéines amyloïdes solubles mal repliées en fibrilles insolubles bien organisées est la marque de plusieurs maladies neurodégénératives telles que les maladies d'Alzheimer, de Parkinson et de Huntington¹.

Dans les maladies d'Alzheimer, les petits oligomères de la protéine A β (dimères à dodécamères) qui précèdent la nucléation des fibrilles sont les espèces les plus toxiques. En raison des difficultés expérimentales, les méthodes de calcul sont un outil essentiel pour obtenir des informations structurales sur les entités toxiques² et les étapes élémentaires du processus d'agrégation. Des modèles simplifiés à gros grains avec des solvants implicites sont couramment utilisés dans ce but³. Cependant, cette stratégie néglige le rôle des corrélations induites par les solvants, un élément essentiel pour la quantification de la cinétique d'agrégation et des durées de vie des états métastables⁴.

Dans ce grand challenge, nous avons effectué des simulations de l'agrégation amyloïde qui n'avaient encore jamais été réalisées. Nous avons considéré une solution d'un millier de peptides amyloïdes A β_{16-22} modélisés par le champ de force gros-grain OPEP (*Optimized Potential for Efficient protein structure Prediction*)⁵ à une résolution quasi atomique et une représentation implicite du solvant. Afin de traiter les corrélations induites par le solvant, y compris donc les interactions hydrodynamiques (HI), le mouvement moléculaire a été couplé à une description efficace de la cinétique des fluides. Ceci a été réalisé par la mise en œuvre d'un nouveau schéma de simulation, le *Lattice Boltzmann Molecular Dynamics*⁴. Contrairement à d'autres méthodes, comme la dynamique brownienne avec HI, notre technique basée sur la méthode *Lattice Boltzmann* locale a une très bonne extensibilité en fonction de la taille du système et est adaptée à un traitement massivement parallèle. Un total de 12 systèmes représentant trois concentrations différentes de la solution (30 mM, 40 mM et 60 mM)

approchant la limite supérieure expérimentale de 20 mM, ont été simulées. Chaque exécution a été faite sur 1024 cœurs de calcul. La composante moléculaire était constituée de 48 000 particules à gros grains et la dynamique des fluides a été traitée sur un réseau de 150^3 points de grille (N_g^3).

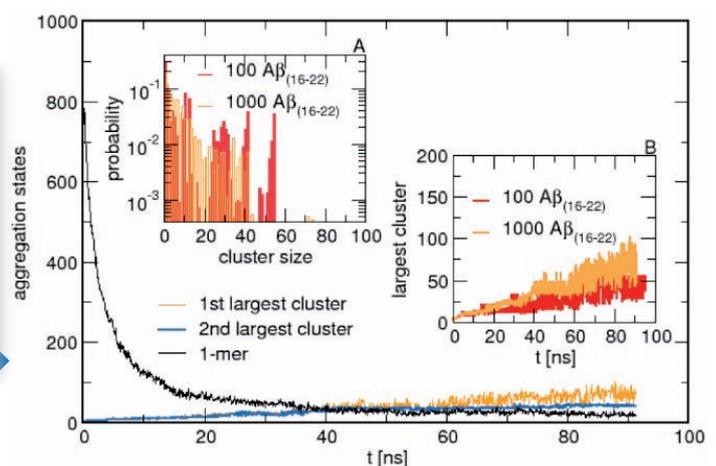
Dans ce traitement, nous avons privilégié une stratégie conservatrice pour la résolution du fluide, le réseau spatial étant $\Delta x = 2 \text{ \AA}$. En modifiant la résolution spatiale du réseau (en augmentant Δx), on peut facilement accroître les performances, le réseau LB croissant en fait en N_g^3 . Si on le transposait à l'échelle atomique, la taille du système serait de 2,4 millions de particules.

Résultats

Dans cette section, nous présentons l'analyse préliminaire de l'une des simulations effectuées au cours de ce grand challenge, correspondant à une concentration de la solution de $c = 60 \text{ mM}$. La propension à l'agrégation des peptides A β_{16-22} est contrôlée par une analyse de l'agrégat en fonction d'un critère de distance entre les centres de masse des peptides. Une distance de coupure $d_c = 12 \text{ \AA}$ a été fixée pour les états avec et sans contact. Le nombre des amas diminue dans le temps selon une décroissance exponentielle étirée $n(t) = e^{-(t/\tau)^\gamma}$ avec pour paramètres $\tau = 8,9 \text{ ns}$ et $\gamma = 0,385$. L'accroissement en exponentielle étirée du processus d'agrégation est probablement la conséquence d'une cinétique avec de multiples temps caractéristiques pour la dynamique des peptides, à la fois ceux absorbés et ceux s'échappant des structures en formation. Le temps moyen obtenu pour l'agrégation est donné par $\langle \tau \rangle = \Gamma(1/\gamma) * \tau/\gamma = 33 \text{ ns}$.

Une compréhension plus précise est obtenue en considérant la façon dont les tailles des différents amas évoluent dans le temps. Plusieurs caractéristiques sont mises en évidence dans la Figure 1, sur laquelle l'évolution temporelle des tailles des deux amas les plus grands est rapportée au nombre de monomères libres en solution. À la fin de la simulation, environ 20 % des peptides du système se retrouvent dans les deux plus grands amas. Les 80 % restants forment des agrégats de tailles plus petites. La poly-dispersion de la solution est bien représentée en considérant la distribution de probabilité des tailles des amas obtenue au cours de la dernière partie de la simulation (encadré A de la Figure 1).

Figure 1 - Évolution dans le temps de la taille du plus grand et du deuxième plus grand agrégat (courbes orange et bleu) et du nombre de monomères libres dans la solution (courbe noire). Encadré A : Distribution des probabilités des tailles cumulées des amas durant les 10 dernières ns des simulations. L'histogramme orange correspond à un système à 1000 A β_{16-22} qui est comparé aux données provenant d'un système plus petit de seulement 100 peptides (histogramme rouge). Encadré B : Évolution dans le temps de la croissance de l'amas le plus grand (en orange) du système à 1000 A β_{16-22} par rapport à celle calculée pour le système plus petit à 100 peptides (en rouge).



Ceci est à comparer aux résultats obtenus pour un système plus petit de seulement 100 peptides. Pour le système simulé durant le grand challenge, l'ensemble des états intermédiaires entre les extrêmes (des monomères libres aux amas les plus grands) sont représentés continuellement, tandis que pour le petit système considéré pendant le même laps de temps, le système se regroupe en quelques structures globales qui contiennent environ 80 % des peptides. L'amas le plus grand croît en permanence durant les simulations, et cette croissance va de pair avec de grandes fluctuations de taille (voir l'encadré B de la Figure 1).

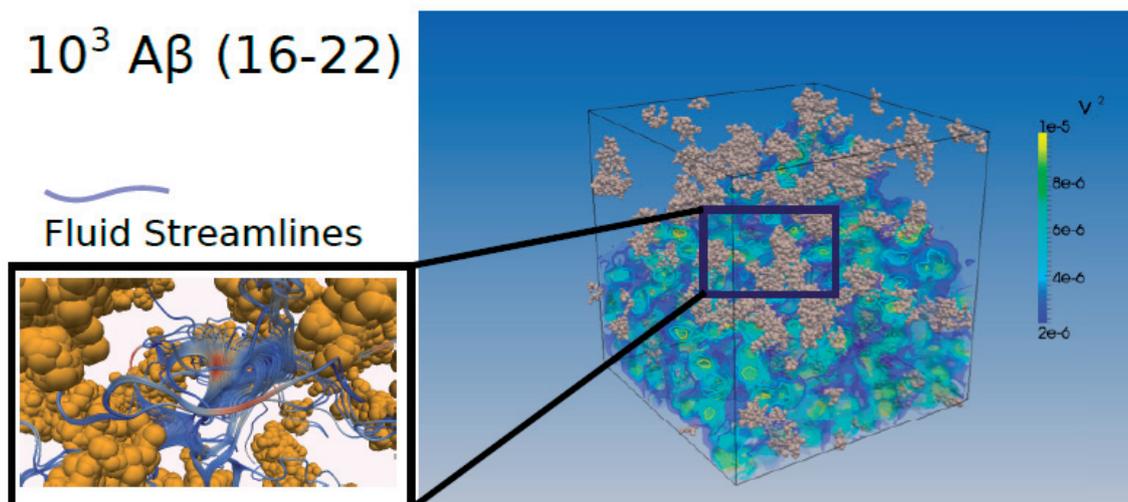
Chacun des agrégats formés, en fluctuant de taille et de forme, produit un flux hydrodynamique qui entraîne la mobilité des entités voisines plus petites. En conséquence, il apparaît clairement que, étant donné la grande taille du système considéré, de nombreuses structures intermédiaires apparaissent et se font concurrence pour attirer les petits oligomères. Ce scénario peut être considéré comme un exemple intéressant de distribution moléculaire hautement concentrée de particules actives qui contribuent toutes à des flux hydrodynamiques⁶.

Ce concept est visualisé sur la Figure 2 où nous avons mis en évidence les lignes de courant de fluide reliant les agrégats de peptides de différentes tailles.

Perspectives

Ce grand challenge nous a donné la possibilité unique de faire progresser un travail sur un système de très grande taille et d'éprouver notre méthodologie jusqu'à présent limitée aux petits systèmes. L'analyse des simulations permettra de caractériser avec de bonnes statistiques les premières étapes de l'agrégation amyloïde pour différentes concentrations. Plus important encore, nous allons maintenant pouvoir quantifier la contribution de l'hydrodynamique à l'association de peptides, de courtes simulations *ad hoc* pouvant être relancées à partir de points de contrôle sélectionnés afin de quantifier la force hydrodynamique s'exerçant sur les monomères ou oligomères lors de la cinétique d'agrégation. Le potentiel de la méthode sera utilisé à l'avenir pour étudier l'agrégation amyloïde dans des conditions analogues à celles des cellules, en incluant le corécepteur principal du cerveau, le prion et les protéines du sérum-albumine humaine.

Figure 2 - Instantané extrait d'une simulation d'un grand système composé de 10^3 monomères du peptide amyloïde $A\beta_{16-22}$. L'iso-vitesse du fluide est représentée par des surfaces colorées. La région agrandie du système met en évidence les courants de fluide cohérents reliant les agrégats amyloïdes.



*À qui la correspondance doit être adressée. Courriel : sterpone@ibpc.fr

¹T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, *Physics Today* 68, 36 (2015).

²A. J. Doig and P. Derreumaux, *Curr. Opin. Struct. Biol.* 30, 50 (2015).

³J. Nasica-Labouze, P.H. Nguyen, F. Sterpone, O. Berthoumieu, N.-V. Buchete, S. Coté, A. D. Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang, and P. Derreumaux, *Chem. Rev.* 115, 3518 (2015).

⁴F. Sterpone, P. Derreumaux, and S. Melchionna, *J. Chem. Theory Comput.* 11, 1843 (2015).

⁵F. Sterpone, S. Melchionna, P. Tuffery, S. Pasquali, N. Mousseau, T. Cragolini, Y. Chebaro, J.-F. St-Pierre, M. Kalimeri, A. Barducci, Y. Laurin, A. Tek, M. Baaden, P. H. Nguyen, and P. Derreumaux, *Chem. Soc. Rev.* 43, 4871 (2014).

⁶A. Mikhailov and R. Kapral, *Proc. Natl. Acad. Sci. USA* 112, E3639 (2015).

Solubilité de CO₂ dans les carbonates fondus

1 Chimie ParisTech, PSL Research University, CNRS, Institut de Recherche de Chimie Paris

2 École Normale Supérieure, PSL Research University, Sorbonne Universités



François-Xavier Coudert¹



Dario Corradini²



Rodolphe Vuilleumier²



Anne Boutin²

Contexte scientifique

La séquestration et la valorisation du dioxyde de carbone constitue un défi majeur de la transition énergétique afin de diminuer la dépendance en énergies fossiles. Dans le cadre d'une collaboration entre des équipes théoriques et expérimentales de *PSL Research University* (École normale supérieure et Chimie ParisTech), nous explorons la voie de l'utilisation des carbonates fondus pour la séquestration et la valorisation du dioxyde de carbone de par leur capacité à dissoudre le CO₂. Les carbonates fondus sont utilisés avec succès dans les piles à combustibles et constituent un milieu prometteur et peu coûteux pour l'électroréduction de CO₂.

Néanmoins, il manque dans la littérature scientifique une compréhension moléculaire de ces processus et une image claire de la réactivité de CO₂. Outre les questions fondamentales, ceci constitue un frein à l'optimisation de systèmes pour l'adsorption et l'électroréduction de CO₂.

Le but de ce projet était de répondre à ce défi par l'étude des aspects moléculaire et thermodynamique de la dissolution de CO₂ dans les carbonates fondus, à l'aide de simulations de dynamique moléculaire à grande échelle. Pour décrire ce système complexe et chimiquement réactif, l'utilisation de méthodes dites « premiers principes », très coûteuses en temps de calcul et en mémoire, est nécessaire.

Nous avons abordé les questions de la structure de CO₂ dans les carbonates fondus, sa solvation par le milieu carbonate, sa réactivité et enfin son transport. Nous avons également ouvert des perspectives par une étude préliminaire de la solubilisation d'autres molécules et d'autres ions, et notamment les espèces CO et O²⁻.

Méthodologie

Les simulations de dynamique moléculaire « premiers principes » reposent sur l'utilisation des outils de la chimie quantique pour traiter les interactions interatomiques au sein du système. Ainsi, les forces interatomiques à chaque instant de la simulation sont obtenues par la méthode de la fonctionnelle de la densité (*Density Functional Theory*, ou DFT) et la dynamique du système est obtenue par intégration numérique des positions des atomes. Ces simulations ont été réalisées avec le logiciel CP2K (<http://www.cp2k.org>), un logiciel *open source* (sous licence GPL) de chimie quantique et de physique de la matière condensée développé de manière collaborative. Le code CP2K, et notamment son module Quickstep de dynamique moléculaire basée sur la DFT, est parallélisé de manière efficace en mode hybride OpenMP + MPI. Cela a permis de tirer pleinement profit du grand nombre de cœurs disponibles sur la machine Turing.

Résultats

Nous avons pu mettre en évidence la formation d'une espèce qui avait été avancée comme une hypothèse par la littérature scientifique antérieure mais dont l'existence n'avait pas été prouvée : l'anion pyrocarbonate C₂O₅²⁻. Espèce transitoire d'une durée de vie de 280 fs dans CaCO₃ fondu à 1773 K, le pyrocarbonate est en équilibre avec CO₂ (Figure 1) et sa formation explique la grande solubilité observée expérimentalement pour le dioxyde de carbone dans les carbonates fondus.

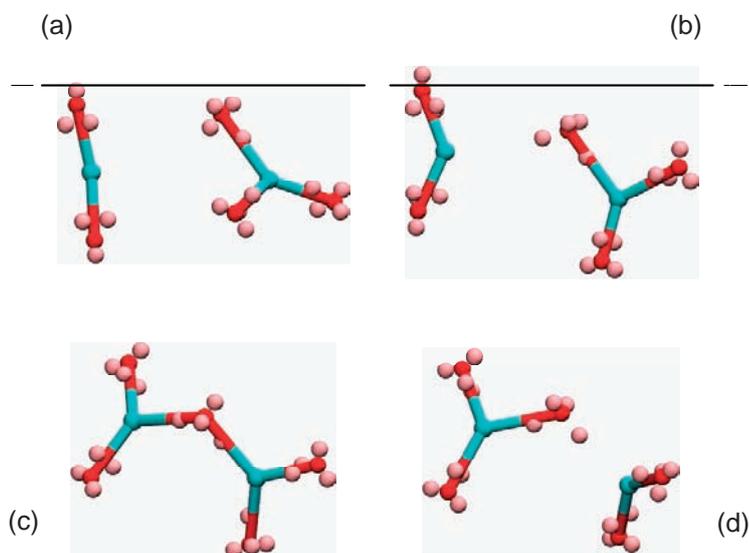


Figure 1 - (a)–(b)–(c) Formation de pyrocarbonate (C₂O₅²⁻) à partir de CO₂ et C₂O₅²⁻. (c)–(d) Dissociation du pyrocarbonate pour reformer CO₃²⁻ et CO₂.

Du point de vue du transport de CO_2 , nous avons pu établir l'existence d'un mécanisme de transport par « saut » d'un atome d'oxygène lors de la formation et dissociation de l'ion pyrocarbonate $\text{C}_2\text{O}_5^{2-}$, conduisant à une diffusion plus rapide que la simple diffusion moléculaire (Figure 2). Nous avons également mis en évidence l'influence de la température sur le transport, grâce à des études à plusieurs températures.

Les résultats obtenus dans le cadre de ce projet sur l'équilibre de formation de l'ion pyrocarbonate (Figure 3) sont inclus dans un article publié dans *Nature Chemistry* (1).

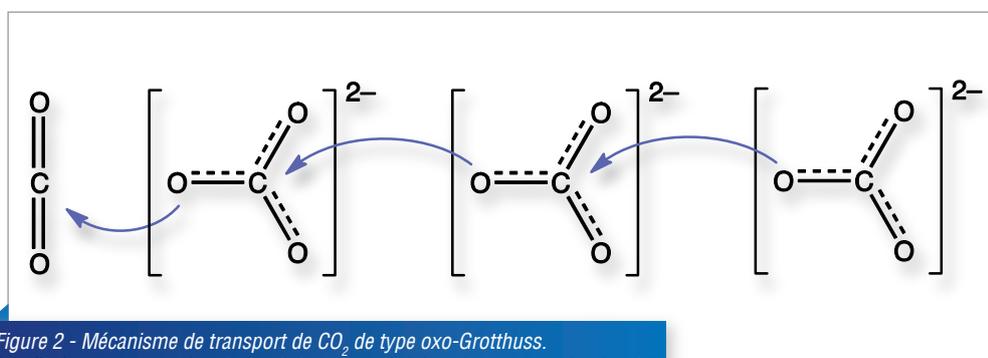


Figure 2 - Mécanisme de transport de CO_2 de type oxo-Grothuss. Les flèches correspondent à des événements disjoints (et non pas simultanés).

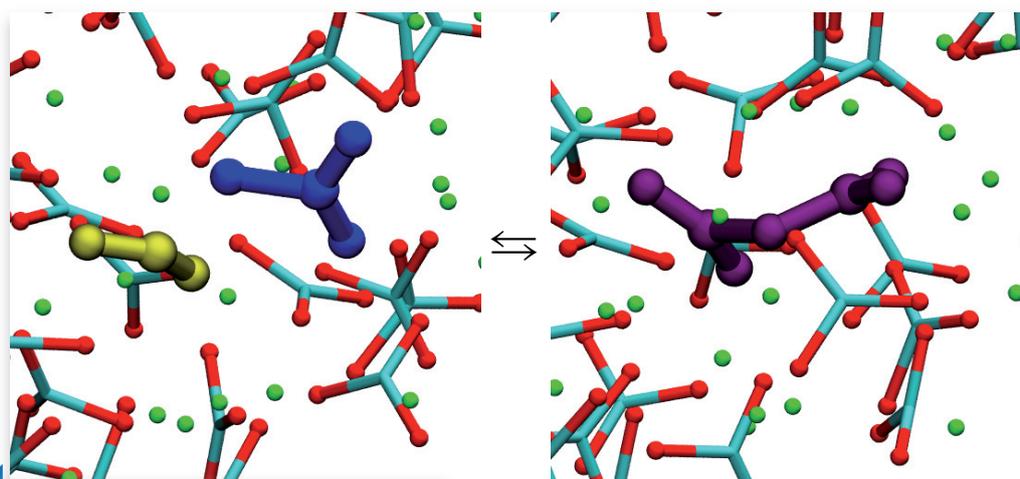


Figure 3 - Équilibre entre CO_2 et l'ion pyrocarbonate

(1) "Carbon dioxide transport in molten calcium carbonate occurs through an oxo-Grothuss mechanism via a pyrocarbonate anion", D. Corradini, F.-X. Coudert and R. Vuilleumier, *Nature Chem.*, 2016, in press, DOI: 10.1038/NCHEM.2450. Lien vers l'article : <http://dx.doi.org/10.1038/nchem.2450>

Centres d'excellence européens

Dans le 8^e plan cadre européen, dit Horizon 2020, la commission européenne a organisé la structuration de son soutien au calcul de haute performance et au traitement des grandes masses de données autour de trois piliers : un pilier technologique, avec le projet ETP4HPC qui regroupe aujourd'hui 72 organisations, un pilier infrastructures avec l'infrastructure et les projets PRACE auxquels l'IDRIS participe depuis leur origine en 2008 et un nouveau pilier scientifique constitué de centres d'excellence qui regroupent chacun un ensemble de partenaires.

Huit projets ont été sélectionnés dans ce cadre en 2015¹, couvrant des domaines scientifiques majeurs comme les énergies renouvelables, la conception de nouveaux matériaux, la chimie atomique et moléculaire, le changement climatique et les outils informatiques pour accroître les performances des applications HPC. Sur ces huit projets, l'IDRIS est impliqué dans deux d'entre eux, renforçant ainsi son engagement européen : EoCoE (*Energy oriented Centre of Excellence for computer applications*) dans le domaine de l'énergie, qui réunit 22 partenaires de 8 pays sous le pilotage de la Maison de la simulation, et E-CAM (*An e-infrastructure for software, training and consultancy in simulation and modelling*), qui réunit 17 partenaires de 10 pays autour de l'organisation du CECAM (Centre européen pour le calcul atomique et moléculaire), dans le domaine de la physico-chimie et des matériaux.

Dans chacun de ces projets, l'IDRIS est impliqué dans les domaines d'activité technologique (dans le pilier technologique transverse du projet EoCoE pour l'utilisation des processeurs *manycore* et des accélérateurs, les entrées-sorties parallèles, l'intégration continue, la visualisation ; dans le projet E-CAM pour le développement de modules logiciels).



Figure 1 - Réunion de lancement du projet EoCoE à Saclay du 19 au 21 octobre 2015



Figure 2 - Réunion de lancement du projet E-CAM à Dublin les 5 et 6 octobre 2015

¹<https://ec.europa.eu/programmes/horizon2020/en/news/eight-new-centres-excellence-computing-applications>

Installation d'un prototype OpenPOWER à l'IDRIS

Dans un contexte d'évolution rapide du HPC et de bouleversements technologiques relatifs aux futures architectures exaflopiques, GENCI, le CEA, le CNRS, l'INRIA et les universités ont décidé de collaborer dans une « Cellule de veille technologique » qui, en regroupant des experts de chaque organisme, vise à mutualiser et mieux tirer parti de l'expérience collective dans ce domaine. Après une première analyse, deux architectures innovantes et à fort potentiel (Intel « Manycore » KNL et accélérée OpenPOWER) ont été sélectionnées pour une étude approfondie via l'acquisition de prototypes. Tandis que la société Atos en a installé un basé sur l'architecture d'Intel au CINES, la machine Ouessant, première phase du prototype OpenPOWER (openpowerfoundation.org), fruit d'une collaboration entre IBM (processeur POWER), NVIDIA (accélérateur) et Mellanox (réseau InfiniBand), l'a été à l'IDRIS.

Elle est composée de trois nœuds de calcul IBM Firestone *bi-sockets* Power8 et d'un nœud de « compatibilité » x86. Les tests viseront aussi bien à quantifier le travail nécessaire au portage des applicatifs sur architecture accélérée que les performances obtenues. Les modèles de programmation à base de directives (OpenMP 4.n ou OpenACC), seront privilégiés et leur maturité et performance seront évaluées. Dans un premier temps, cette plateforme est ouverte aux porteurs de la vingtaine de codes sélectionnés, représentatifs des différents domaines scientifiques. Plus tard, elle sera ouverte à toutes les communautés désirant faire des tests. Une deuxième phase proposera 12 nœuds de calcul IBM Garrison comprenant chacun deux processeurs Power8+, des connexions NVLink1 vers quatre cartes accélératrices NVIDIA Pascal, l'interface CAPI et un réseau InfiniBand Mellanox EDR.

Conférence Supercomputing 2015

La 27^e conférence Supercomputing, rendez-vous incontournable pour les acteurs industriels et académiques du calcul intensif et du traitement des grandes masses de données, s'est tenue à Austin du 15 au 20 novembre 2015, regroupant près de 12 000 participants.

Pour la deuxième année consécutive, le CC-IN2P3 et l'IDRIS ont représenté conjointement le CNRS sur la partie salon de la conférence. Les deux centres informatiques du CNRS ont ainsi pu démontrer à nouveau leur complémentarité : si l'IDRIS se concentre plus particulièrement sur le calcul haute performance (HPC), le CC-IN2P3 est lui davantage orienté vers la gestion et le traitement des données, en particulier dans les domaines de la physique des hautes énergies et des astroparticules. GENCI (Grand équipement national de calcul intensif) était également présent sur le stand aux côtés de l'IDRIS.

Ce stand a aussi été l'occasion de présenter deux projets plus en détail, d'une part l'infrastructure mise en place par le CC-IN2P3 pour l'expérience LSST (*Large Synoptic Survey Telescope*), basée sur les données qui seront produites par un nouveau télescope installé au Chili, dont les observations devraient permettre de mieux comprendre la nature de l'énergie noire, responsable de l'accélération de l'expansion de l'univers. Pendant les dix années du projet, 37 milliards d'objets seront stockés, pour un volume total de 500 pétaoctets. De son côté, l'IDRIS détaillait le projet E-Biothon,



une plate-forme qu'il héberge, issue de la collaboration entre le CNRS, IBM, INRIA, l'Institut français de bioinformatique et SysFera, destinée à accélérer et faire progresser la recherche biomédicale.

Consommation des heures de calcul en 2015 et attributions en 2016

Voici le bilan des heures consommées en 2015 ainsi que celui des heures attribuées pour la première session de la campagne DARI de l'année 2016 sur Turing et Ada.

Bilan des heures consommées en 2015

Turing

La disponibilité de Turing en 2015 (en dehors des maintenances programmées) a été de 99,68 %. Au total, 577 millions d'heures auront été utilisées par 100 projets DARI des différents comités thématiques.

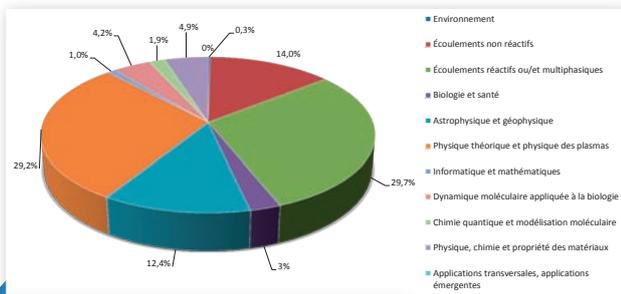


Figure 1 – Répartition par comités thématiques des heures de calcul consommées (en pourcentages) en 2015 sur la machine Turing

Comme on peut le voir sur la figure 1, les comités dont les projets ont été les plus consommateurs sont le CT2b (écoulements réactifs ou/et multiphasiques) et le CT5 (physique théorique et physique des plasmas).

Ada

Sa disponibilité (hors maintenances programmées) en 2015 a été de 99,41 %. Plus de 65 millions d'heures de calcul ont été consommées par 242 projets.

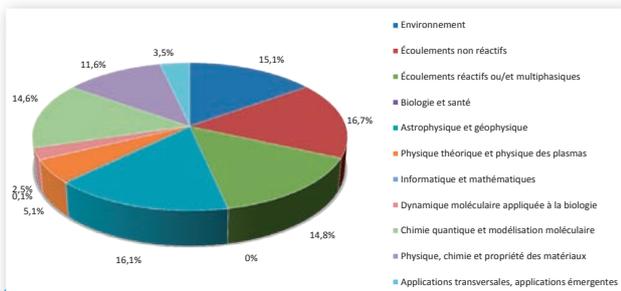


Figure 2 – Répartition par comités thématiques des heures de calcul consommées (en pourcentages) en 2015 sur la machine Ada

On peut constater sur la figure 2 que les comités dont les projets ont été les plus consommateurs sont le CT2a (écoulements non réactifs), le CT4 (astrophysique et géophysique), le CT1 (environnement) et le CT2b (écoulements réactifs ou/et multiphasiques).



Turing Blue Gene/Q

De plus, ce sont 22 dossiers d'accès préparatoires (15 sur Ada et 7 sur Turing) qui ont été acceptés en 2015, allouant des ressources (respectivement 225 000 h sur Ada et 350 000 h sur Turing) pour favoriser le portage et l'optimisation des applications sur ces calculateurs.

Bilan des heures attribuées lors de la première session DARI 2016

Pour cette première session de la campagne 2016, plus de 626 millions d'heures ont été attribuées sur Turing à 99 projets des différents comités thématiques. Concernant Ada, ce sont plus de 66 millions d'heures qui ont été attribuées à 218 projets. La répartition des heures de calculs attribuées par comités thématiques lors de cette 1^{re} session est présentée dans les figures 3 et 4.

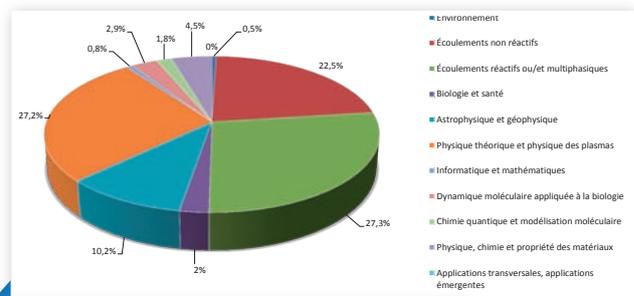


Figure 3 – Répartition par comités thématiques des heures de calcul attribuées (en pourcentages) lors de la 1^{re} session 2016 sur la machine Turing

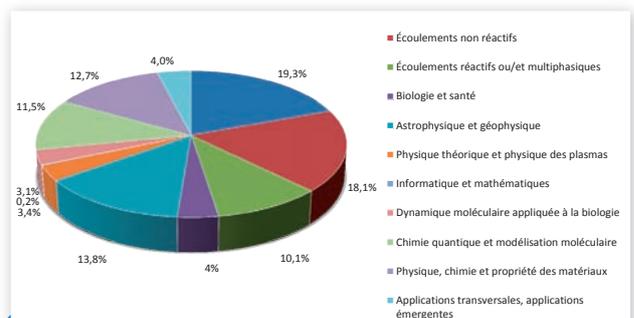


Figure 4 – Répartition par comités thématiques des heures de calcul attribuées (en pourcentages) lors de la 1^{re} session 2016 sur la machine Ada

Bilan des actions de formation 2015

Les activités de formation sont pour l'IDRIS une composante essentielle de son rôle de pôle de compétences en calcul de haute performance.

Une dizaine d'ingénieurs de l'équipe Support aux utilisateurs donnent des formations principalement orientées vers les méthodes de programmation parallèles (MPI, OpenMP, programmation hybride MPI/OpenMP) et les langages de programmation scientifiques généraux (C, Fortran 95 et 2003). Ces formations sont dispensées à l'IDRIS ou à l'extérieur et sont accessibles à toute la communauté académique et industrielle utilisatrice du calcul intensif.

Tous les supports de cours IDRIS et les travaux pratiques associés sont disponibles librement sur le site de l'IDRIS : www.idris.fr/formations/supports_de_cours.html

L'IDRIS participe également aux formations PATC (*Prace Advanced Training Center*) françaises coordonnées par la Maison de la simulation. En 2015, les ingénieurs de l'IDRIS ont ainsi animé 2 sessions de formation PATC concernant la programmation hybride et les entrées-sorties parallèles.

En 2015, l'IDRIS a dispensé au total 19 sessions de formation (dont 3 à l'extérieur), qui ont été suivies par 260 personnes (214 à l'IDRIS, 46 à l'extérieur) pour un total de 62,5 journées de formation.

Par ailleurs, les salles de conférence, formation et réunion de l'IDRIS peuvent être mises à disposition, sous certaines conditions, à des organismes relevant du campus Paris-Saclay et plus généralement de la région Île-de-France. En 2015, nous avons ainsi accueilli des formations internes de membres du personnel du laboratoire IPSL, du service financier et comptable de la délégation régionale Île-de-France Sud du CNRS et deux tutoriels CECAM (Centre européen de calcul atomique et moléculaire), soit en tout plus de 150 personnes. Pour tout renseignement complémentaire, contactez la communication de l'IDRIS (communication@idris.fr).

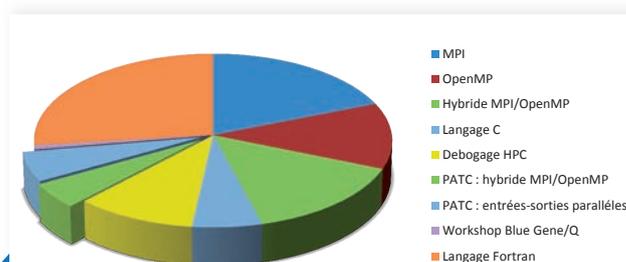


Figure 1 – Bilan des actions de formation 2015 (en nombre d'inscrits)

CALENDRIER DES FORMATIONS IDRIS PROGRAMMÉES AUX 2^e et 3^e TRIMESTRES 2016

TITRE DE LA FORMATION	DATE DE DÉBUT	DURÉE
Programmation hybride MPI/OpenMP	10/05/2016	2 jours
Langage C	30/05/2016	5 jours
MPI / OpenMP	06/06/2016	5 jours
Fortran avancé	14/06/2016	4 jours
Portage et optimisation sur Blue Gene/Q	29/06/2016	3 jours
Fortran base	20/09/2016	3 jours
MPI	26/09/2016	4 jours

Ces dates sont susceptibles d'être mises à jour. N'hésitez pas à consulter le serveur Web des cours de l'IDRIS : <https://cours.idris.fr> Vous y trouverez le catalogue des formations et les modalités d'inscription.

Les formations IDRIS sont gratuites pour le personnel du CNRS ou des universités. Elles sont aussi accessibles aux entreprises publiques ou privées via CNRS Formation Entreprises : cnrsformation.cnrs.fr

