

Des supercalculateurs pour mieux connaître l'Univers

La diffusion du parallélisme dans toute la pyramide de l'informatique et la prise en compte de contraintes multiples, comme la gestion des données ou l'architecture hiérarchisée des supercalculateurs, plongent les spécialistes du calcul scientifique au cœur d'une tourmente. En effet, l'installation, l'utilisation et l'administration de leurs nouveaux moyens de calcul s'opèrent désormais à l'instar des très grands équipements et la communauté des usagers doit donc apprendre à s'organiser en équipes pluridisciplinaires autour d'eux.



Supercalculateur BULL hybride installé au CCRT en 2009.

CEA

Ainsi, pour réussir pleinement les changements d'échelles résultant des avancées technologiques enregistrées dans le domaine du calcul de haute performance (le HPC pour *High Performance Computing*), les chercheurs en sciences fondamentales et appliquées doivent avoir accès à des ressources de calcul pouvant aller jusqu'au calculateur entier, au moins pendant des périodes limitées de temps. Il existe un moment propice pour une utilisation dite « grand challenge ». Il s'agit de la période de mise en production de la machine qui va de 3 à 6 mois après la réception du calculateur. Et pour cause, il s'agit d'un moment privilégié pendant lequel les équipes du centre de calcul (ingénieurs systèmes et spécialistes applicatifs) et les experts du constructeur informatique se retrouvent tous ensemble et mobilisés sur le site pour régler d'éventuels problèmes liés au démarrage de l'instrument. Les chercheurs peuvent alors bénéficier de cette étroite collaboration entre spécialistes pour optimiser leur logiciel de simulation et pour espérer franchir de nouvelles étapes dans la réalisation de simulations de très grande taille. Ainsi, les simulations de quelques milliers de processeurs qui relevaient, il y a peu de temps encore, du défi technique, sont-elles désormais très fréquentes.

En astrophysique, les grands challenges visent une meilleure prise en compte des couplages d'échelles inhérents à la plupart des phénomènes physiques rencontrés, cela en augmentant la **résolution spatiale**

et/ou temporelle. La très grande quantité de mémoire disponible sur les supercalculateurs, associée à des techniques numériques novatrices, favorise l'accès à des résolutions spatiales de plus en plus importantes. L'objectif reste de s'approcher des échelles de dissipation dans les écoulements **turbulents**, de résoudre les cœurs protostellaires lors de la simulation d'un nuage moléculaire, ou encore de simuler finement les **galaxies** dans un contexte cosmologique. De cette dernière problématique naquit, en 2007, le code RAMSES, développé par le SAp, pour étudier la formation des grandes structures et des galaxies. Il s'agit d'un code s'inscrivant dans le cadre du projet Horizon visant à fédérer les activités de simulation numérique autour d'un projet ciblé sur l'étude de la formation des galaxies. Cette réflexion fut menée pendant la période de démarrage du calculateur Bull Platine, au Centre de calcul recherche et technologie (CCRT) du CEA, avec pour objectif la simulation de la formation d'une moitié de l'Univers observable. Pour la première fois dans l'histoire du calcul scientifique, il fut possible de décrire une galaxie comme la **Voie lactée** avec plus d'une centaine de particules tout en couvrant la moitié de l'Univers observable. Pour simuler un tel volume avec autant de détails, les acteurs du projet Horizon utilisèrent 6 144 processeurs Intel Itanium2® du calculateur Bull Platine pour activer le programme RAMSES à plein régime. Ce logiciel de simulation met en jeu une grille adaptative permettant d'atteindre une finesse spatiale inégalée. Avec près de 70 milliards de particules et plus de 140 milliards de mailles, ce grand challenge représente le record absolu pour un système à N corps modélisés par ordinateur.

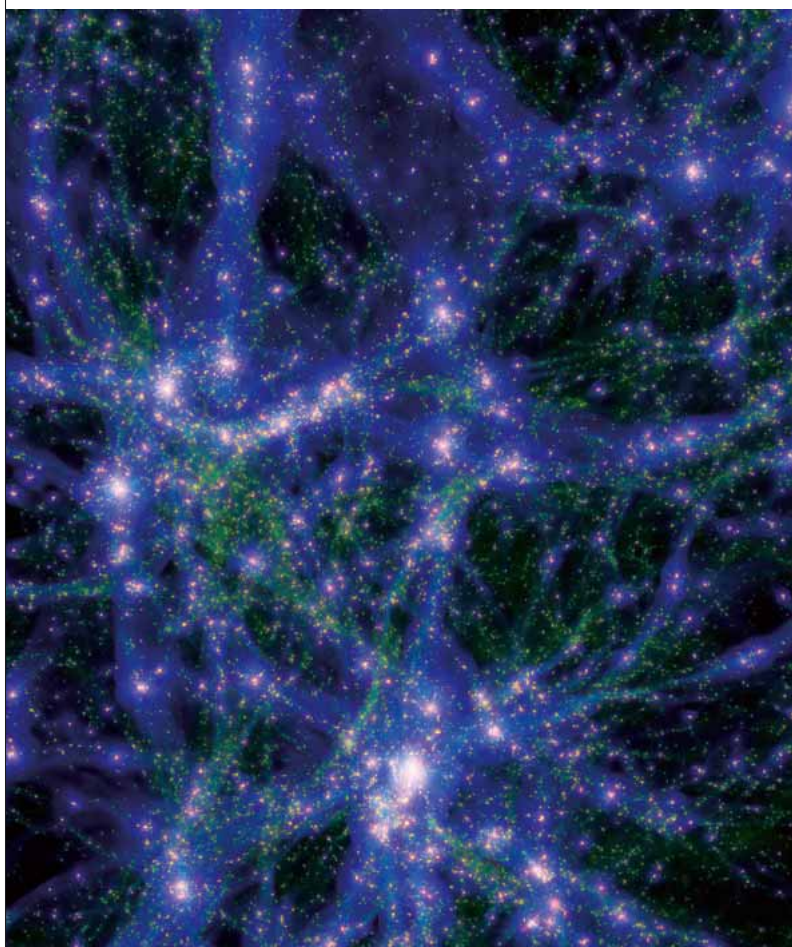
Cet exemple montre bien, qu'en matière de simulation, les avancées promises par les ordinateurs de grande puissance ne s'obtiennent qu'au prix d'une maîtrise de la complexité – celle concernant à la fois les modèles physiques, les méthodologies et les **algorithmes** numériques, les méthodologies et les techniques de programmation et d'optimisation parallèles. De plus, la nécessité de contenir la consommation électrique a favorisé l'apparition d'un nouveau type de supercalculateurs capables de combiner un très grand nombre de processeurs généralistes avec des processeurs spécialisés (processeurs graphiques, reconfigurables, vectoriels...). Dès l'été 2009, la communauté scientifique française a pu accéder à ces supercalculateurs hybrides grâce à la machine BULL installée au CCRT. Avec plus de 2 100 processeurs Intel de nouvelle génération, associés à 48 serveurs graphiques Nvidia, c'est le premier supercalculateur de ce type implanté en Europe.

> Pierre Leca et Christine Menaché

Département des sciences de la simulation
et de l'information (DSSI)
Direction des applications militaires (DAM)
CEA Centre DAM Ile-de-France

> Édouard Audit

Service d'astrophysique (SAp)
Institut de recherche sur les lois fondamentales de l'Univers (Irfu)
Direction des sciences de la matière (DSM)
Unité mixte de recherche astrophysique
interactions multi-échelles
(CEA-Université Paris 7-CNRS)
CEA Centre de Saclay (Orme des Merisiers)



Le code RAMSES permet l'étude des structures de l'Univers à grande échelle ainsi que la formation des galaxies. La plus grande simulation de la formation de ces structures a été réalisée dans le cadre du projet Horizon, soutenu par l'Agence nationale pour la recherche (ANR).

La visualisation des simulations astrophysiques

La **simulation** numérique des **plasmas** astrophysiques produit des données de taille et de complexité considérables. Pour les comprendre et les interpréter, les astrophysiciens doivent les visualiser avec des outils logiciels adaptés. À cette fin, le logiciel SDvision (*Saclay Data Visualisation*), développé à l'Irfu, permet une visualisation interactive de simulations produites sur des ordinateurs massivement parallèles, utilisant des milliers de processeurs, produisant des données de taille pouvant dépasser le **téraoctet** (figure 1). La complexité des simulations réside essentiellement dans la mise en œuvre du maillage de l'espace dont la résolution varie en fonction de différents critères physiques.

La technique numérique du maillage à résolution adaptative (figure 2) concentre le temps de calcul, et la mémoire allouée, sur des zones accueillant des phénomènes significatifs comme la formation d'**amas de galaxies**. En revanche, dans les zones à plus faible densité, les calculs restent à des résolutions plus modestes. Pour visualiser et analyser de telles structures de données, les chercheurs ont dû recourir à des **algorithmes** spécifiques et relever un défi : celui de la gestion de la mémoire. Il leur a fallu, en outre, s'adapter à divers modes d'utilisation. La visualisation et l'analyse de telles structures de données représentent un véritable défi, qui a conduit au développement d'algorithmes spécifiques, notamment en termes de gestion de la mémoire. En outre, le logiciel SDvision est adapté à des modes d'utilisation divers, par exemple sur des plateformes de visualisation locales bénéficiant de l'accélération matérielle des cartes graphiques modernes ou distantes sur des nœuds graphiques proposant d'importantes ressources (multicœurs, mémoire partagée

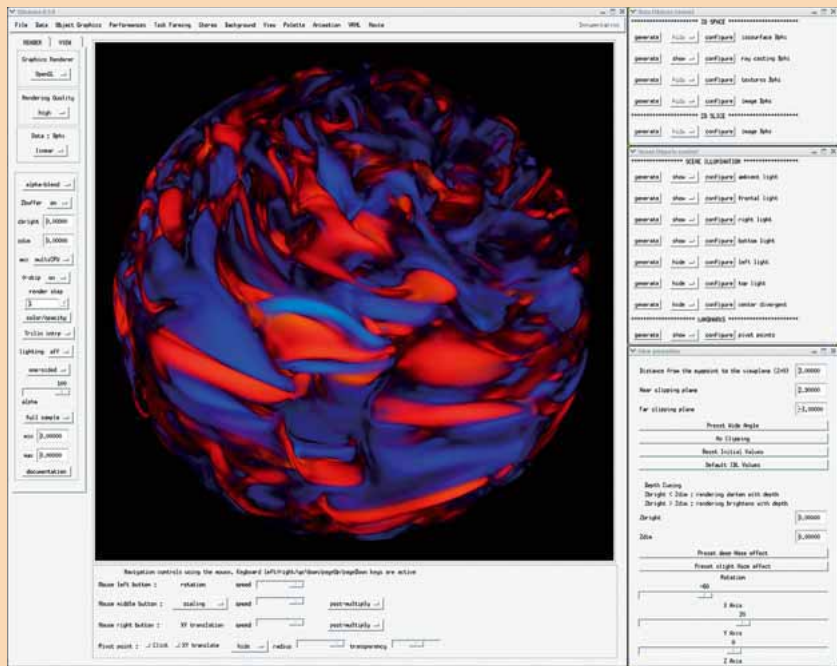


Figure 1. Le logiciel de visualisation interactive SDvision utilisé, ici, pour visualiser le champ magnétique obtenu dans une simulation de la zone convective solaire.

élevée dépassant 100 Gigaoctets par nœud). Véritable instrument d'exploration des simulations, ce logiciel permet de naviguer interactivement et en immersion dans les données. Elles contribuent également à découvrir ou à étudier des structures astrophysiques à échelles multiples : par exemple, la formation des grandes structures cosmologiques, des **galaxies**, **turbulences** du milieu interstellaire participant à la création des « cœurs protostellaires » où se forment les **étoiles**, dynamique des **disques d'accrétion** planétaires, **magnétohydrodynamique** de la zone convective solaire...

Le logiciel est employé dans une salle de visualisation stéréoscopique installée au Service d'astrophysique (SAP/Irfu). Il s'agit d'une technique qui appréhende complètement la structuration des volumes simulés dans les trois dimensions. Ces développements, qui ont pour moteur principal la recherche fondamentale en astrophysique, devraient également bénéficier à d'autres domaines de la recherche et de la technologie utilisant la simulation numérique : par exemple, pour explorer des simulations du transport turbulent dans le plasma du futur réacteur expérimental ITER (pour *International Thermonuclear Experimental Reactor*) actuellement en construction sur le centre du CEA de Cadarache. Il s'agit d'une collaboration avec l'Institut de recherche sur la fusion magnétique (DSM).

La montée en puissance des grands centres de calculs nationaux laisse prévoir l'obtention de simulations toujours plus massives et plus complexes mais aussi de nouveaux défis à relever en matière de visualisation des simulations destinées à l'astrophysique.

➤ **Daniel Pomarède**
et **Bruno Thooris**

Service de l'électronique, des détecteurs
et de l'informatique (Sedi)
Institut de recherche sur les lois fondamentales
de l'Univers (Irfu)
Direction des sciences de la matière (DSM)
CEA Centre de Saclay

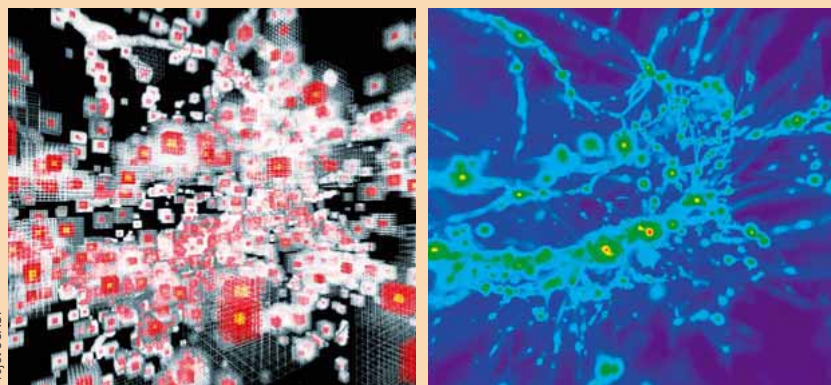


Figure 2. Visualisation d'une simulation cosmologique. À gauche : le maillage à résolution adaptative employé dans la simulation, représenté à travers les mailles de résolutions les plus élevées. À droite : la distribution de la densité de matière calculée sur ce maillage.

Godunov, une plate-forme numérique pour l'enseignement-recherche

Il devient indispensable d'enseigner les techniques et les méthodologies destinées à la **modélisation** et la **simulation numérique** aux étudiants, aux doctorants et postdoctorants, ainsi qu'aux chercheurs et ingénieurs, impliqués ou non en astrophysique, et qui souhaitent voir évoluer leurs connaissances. D'où la mise en place, depuis 2007, de la plate-forme numérique pour l'enseignement et la recherche *Godunov* (du nom du mathématicien qui développa une méthode classique de résolution des équations de l'hydrodynamique). Il s'agit d'une collaboration entre le SAP, l'Unité d'enseignement en sciences de la matière et de simulation (UESMS) de l'Institut national des sciences et techniques nucléaires (INSTN) et l'Université Paris 7/ Denis-Diderot. Située sur le centre du CEA de Saclay, cette plate-forme a été équipée de vingt postes de travail, reliés par un réseau rapide à un *cluster* de calcul, c'est-à-dire un ordinateur

(1) Technique qui permet de cartographier la distribution de la matière noire dans l'Univers et de caractériser les propriétés de l'énergie noire.

parallèle comportant 240 cœurs de calcul. Bien que la plate-forme soit également destinée à la recherche, les activités pédagogiques demeurent néanmoins prioritaires. Ainsi, pendant les sessions de formations, le *cluster* demeure exclusivement réservé aux étudiants afin qu'ils s'initient au calcul haute performance à l'aide d'un outil à la fois performant et disponible, avant de se confronter à des ordinateurs de plus grande taille. Les enseignements correspondent à un niveau postdoctoral, principalement dans le cadre de l'École doctorale d'astrophysique, mais aussi à celui des masters dans lesquels le CEA s'implique *via* l'INSTN : par exemple, astronomie & astrophysique, modélisation et simulation (M2S), matériaux pour les structures et l'énergie (MSE) et interaction climat et environnement (ICE). À ces enseignements s'ajoutent deux formations continues. L'une menée conjointement par l'INSTN et le SAP où sont abordées les méthodes de **Monte-Carlo** et les techniques de programmation parallèle (bibliothèque MPI) ; l'autre étant dédiée aux utilisateurs du Centre de calcul du CEA (CCRT).

Outre sa vocation pédagogique, *Godunov* sert au traitement de données pour l'astrophysique. Par exemple, la préparation des **algorithmes** de traitement de données émanant du satellite GLAST (pour *Gamma-ray Large Area Space Telescope*) et la réalisation du catalogue des sources. D'autres expériences, comme le *weak lensing*⁽¹⁾, bénéficient également de cette plate-forme pour leurs besoins de calcul, importants mais ponctuels, en matière de traitement des images.

> Édouard Audit

Service d'astrophysique (Sap)
Institut de recherche sur les lois fondamentales de l'Univers (Irfu)
Direction des sciences de la matière (DSM)
Unité mixte de recherche astrophysique interactions multi-échelles (CEA-Université de Paris 7-CNRS)
CEA Centre de Saclay (Orme des Merisiers)

> Constantin Meis

Institut national des sciences et techniques du nucléaire (INSTN)
Unité d'enseignement en sciences de la matière et de simulation (UESMS)
CEA Centre de Saclay



Enseignement dans la salle de calcul de la plate-forme Godunov : simulation de plasma.