

CALCULS D'INFRASONS SUR TERA 10

B. DESPRÉS, H. JOURDREN, P.-F. PISERCHIA, O. GAINVILLE
CEA - DAM - Île-de-France

Dans toute discipline technique, il est important de savoir quand un seuil inatteignable auparavant est atteint, et même dépassé. C'est ce qui vient de se passer dans le courant de l'année 2006 pour les calculs tridimensionnels directs d'infrasons dans l'atmosphère terrestre à l'aide du supercalculateur TERA 10. Les infrasons sont des ondes acoustiques de fréquence faible susceptibles de se propager sur de très longues distances dans l'atmosphère terrestre. La mesure physique des infrasons, puis leur analyse et leur calcul, contribue aux moyens de surveillance internationaux.



Qu'est ce qu'un infrason ?

Les ondes infrasonores, ou infrasons, sont des ondes acoustiques qui se propagent par réflexions multiples sur les différentes couches de l'atmosphère et sur le sol. Leur fréquence inférieure à 20 Hz se situe en dessous du domaine audible (*d'où leur nom*). Leur vitesse au sol, de l'ordre de 330 m/s, peut être modulée par le vent et la température en altitude. Les infrasons sont d'origine naturelle ou artificielle. Certaines sources infrasonores sont suffisamment puissantes pour secouer l'atmosphère en faisant plusieurs fois le tour de la Terre. Les stations infrasons du TICE (*Traité d'Interdiction Complète des Essais*) ont pour mission de surveiller les signaux acoustiques dans l'atmosphère pour détecter les éventuels contrevenants à ce traité. Cela explique l'intérêt majeur de tester régulièrement ces stations infrasons avec des sources dont la position est connue avec précision.

Nous pouvons distinguer deux origines d'infrasons :

- ceux issus des phénomènes naturels qui agitent l'atmosphère (*orages*) et le globe terrestre. Par exemple, les météorites qui se désintègrent dans notre atmosphère, les séismes qui ébranlent le sol, génèrent des infrasons ;
- ceux d'origine humaine, créés par les explosions (*accidents, tirs de carrière, essais*).

Comment calculer les infrasons ?

Le calcul numérique des infrasons repose sur la discrétisation d'un système d'équations aux dérivées partielles décrivant le phénomène physique. La difficulté principale provient de la longueur caractéristique du phénomène physique. En effet, il est nécessaire d'échantillonner suffisamment, au moins 10 points de discrétisation par longueur d'onde, pour un calcul direct usuel (*application particulière de la règle de Shannon*). Pour des domaines tridimensionnels de plusieurs centaines de longueurs d'onde par dimension d'espace, comme c'est le cas pour l'atmosphère, cela conduit rapidement à un milliard de mailles.

Des méthodes alternatives existent. Elles consistent en une approximation raisonnable des équations, pour contourner cette règle d'échantillonnage et réduire la charge de calcul. C'est le cas des méthodes de rayons de type *Hamilton-Jacobi*, ou des méthodes paraboliques que les géophysiciens du CEA - DAM exploitent. Cela se fait au prix d'une approximation supplémentaire, bien maîtrisée, grâce à des comparaisons avec des méthodes avancées, comme la résolution directe des équations d'*Euler* linéarisées. Avec la résolution directe, la charge de calcul peut devenir gigantesque, à cause de la règle d'échantillonnage du signal, entraînant une grande difficulté de mise en œuvre. C'est là qu'intervient l'utilité des calculateurs massivement parallèles tels que le calculateur TERA 10. La place mémoire disponible rend possible une discrétisation suffisamment fine de l'atmosphère. Le nombre de processeurs diminue le temps de restitution du calcul, à condition que ceux-ci soient bien utilisés en parallèle.

Description du calcul sur TERA 10

Nous commençons par décrire les conditions générales de calcul. Dès la mise en service de la machine TERA 10, des simulations "en vraie grandeur" ont été lancées pour évaluer la stabilité matérielle et logicielle de la configuration. Il s'agissait d'une période privilégiée pour les équipes de développement, qui leur a permis de vérifier le niveau de portabilité des codes, le bien fondé des dernières évolutions de leurs architectures, tout en engrangeant des premières simulations d'exception.

Avec 5 000 processeurs double-cœur et plusieurs mégawatts de consommation électrique, la machine TERA 10 est à ce jour la deuxième plus puissante en Europe. TERA 10 offre, avec 8 processeurs double-cœur par nœud, (au lieu de 4 processeurs simple-cœur sur TERA 1) une bouffée d'oxygène appréciable pour les codes vectoriels, un facteur 4 sur le nombre maximal de processeurs par calcul, et un facteur 12 sur la mémoire. Ce choix se traduit par une évolution aux conséquences parfois subtiles pour la gestion mémoire, le passage d'une configuration de type grappe de nœuds SMP (TERA 1) à une configuration grappe de nœuds NUMA (TERA 10). Les gains en puissance apportés par le nouveau processeur Itanium (au lieu du processeur Alpha sur TERA 1) sont impressionnants. Un premier très gros calcul mené sur TERA 10 en février 2006 a permis d'en apprécier les enjeux. Le calcul a été effectué avec la plate-forme C++ *Dragst-R* sur plusieurs briques technologiques "calcul haute performance", notamment :

- le mode de rangement mémoire du *benchmark* de recette TERA-TF permettant une utilisation quasi-optimale par directions alternées de la mémoire cache de chaque processeur ;
- la souche récente [1] de schémas numériques d'ordre très élevé pour l'équation des ondes et l'advection, mise au point pour tirer le maximum du processeur Itanium, soit un très grand nombre d'opérations "utiles" pour peu de requêtes mémoire,
- la bibliothèque MPC d'échanges de messages MPI/multithread ;
- la bibliothèque *Hercule* de traitement de très gros volumes de données.

Ce calcul a reçu le label *Grands Challenges*, qui distingue les réalisations remarquables en "calcul parallèle haute performance"

Le calcul principal est une simulation *Euler* 3D de l'expérience *Misty Picture* (explosion de 4 685 tonnes de nitrate d'ammonium et d'essence, le 14 mai 1987, au Nouveau Mexique - USA, à la surface du sol). Le domaine de simulation 3D couvre 4 millions de km² au sol (2 000 x 2 000 km²) et s'étend jusqu'à 200 km d'altitude. Les résultats de la simulation sont résumés sur la figure 1.

Plusieurs coupes du domaine de calcul sont données sur la figure 2. Un schéma d'ordre 17 a été utilisé pour résoudre l'équation des ondes avec vents, ce qui, pour un signal de 0,1 Hz et 8 mailles par longueur d'onde, conduit à une résolution spatiale de 400 mètres au sol, et un total de 10 milliards de mailles.

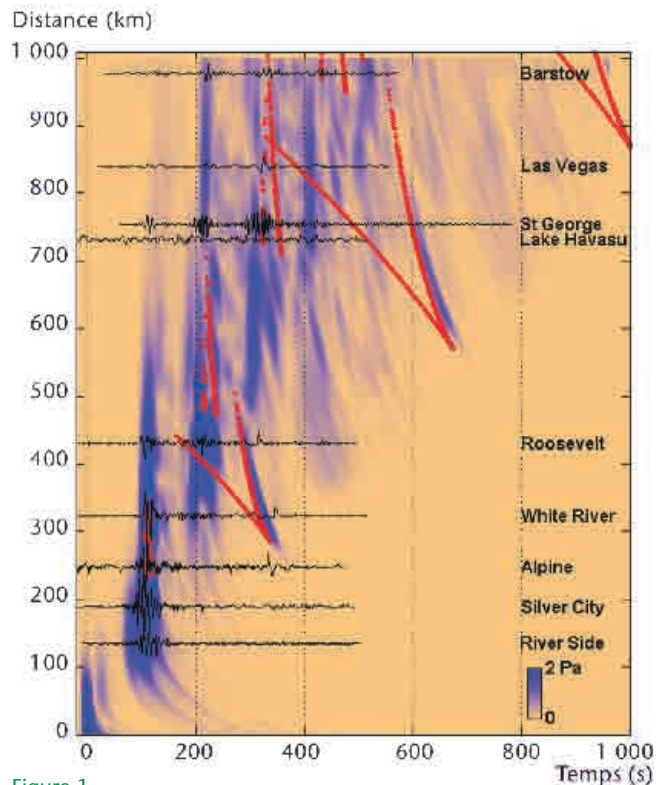


Figure 1

L'énergie de signaux synthétiques (en bleu) est superposée aux mesures (en noir) pour des stations situées à différentes distances de la source. Cette représentation permet d'identifier le trajet des ondes dans l'atmosphère, et d'estimer la pertinence des calculs. Les lignes rouges donnent les temps d'arrivée issus du tracé des rayons.

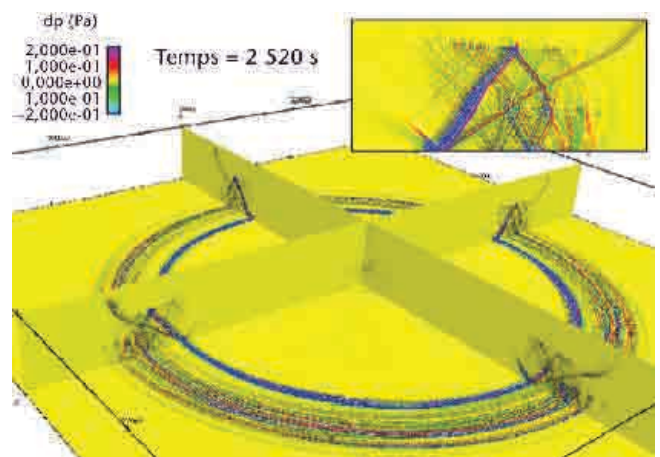


Figure 2

Calcul *Dragstser*.

Sur cette figure, les isolignes de pression sont représentées sur deux plans verticaux et sur un plan horizontal. Les plans se coupent à angle droit. Nous observons que l'onde de pression (infrasons) créée au centre s'est propagée en se réfléchissant à mi-hauteur (vers 100 km d'altitude). Un agrandissement illustre la structure riche de l'onde de pression au sol. Le vent en altitude explique la légère dissymétrie du champ de pression.

Une fraction de la machine TERA 10 a suffi pour mener le calcul jusqu'au temps d'arrivée du signal infrasonore sur les sondes les plus éloignées (3 500 secondes), soit exactement 384 processeurs durant 2 jours et demi, pour une puissance soutenue constatée de 0,85 téraflops, et 2,2 gigaflops par processeur. La même simulation menée durant 10 heures sur 1 936 processeurs a conduit à 5,8 téraflops. Ce chiffre est d'excellent augure pour la configuration complète TERA 10, qui devrait permettre, sur ce type d'application, une puissance soutenue de 20 téraflops.

Conclusions et perspectives

D'une part, les études d'extensibilité ont montré que nous possédons à présent une excellente maîtrise des méthodes de programmation parallèle sur TERA 10, pour des problèmes qui se prêtent favorablement à un traitement parallèle.

D'autre part, les avancées importantes en calcul d'infrasons, qui ont donné lieu à une publication [2] en cours d'évaluation, ainsi qu'à des communications dans des conférences internationales [3], permettent d'ores et déjà d'asseoir la confiance que nous avons dans nos outils de calculs des infrasons.

Références

- [1] S. DEL PINO, H. JOURDREN, "Arbitrary high-order schemes for the linear advection and wave equations: application to hydrodynamics and aeroacoustics", *C. R. Acad. Sci. Paris, Série I* 342(6), p. 441-446 (2006).
- [2] B. DESPRÉS, S. DELPINO, H. JOURDREN, P. HAVÉ, P.-F. PISERCHIA, "3D Finite Volume Simulation of Acoustic Waves in the atmosphere", soumis à *Computer and Fluids*.
- [3] O. GAINVILLE, P.-F. PISERCHIA, P. BLANC-BENON, J. SCOTT, *AIAA n°2006-2451* (2006).

TERA 10

Le supercalculateur TERA 10, un nouvel outil pour la recherche

Conçu, fabriqué et installé par la société Bull, TERA 10 (figure) est un supercalculateur parallèle, c'est-à-dire capable de faire travailler simultanément, sur une même simulation, un grand nombre de processeurs. Il comprend 5 000 processeurs double-cœur Intel de type "Montecito", qui fournissent chacun une puissance de calcul de 12,8 millions d'opérations par seconde.

L'élément de base de TERA 10 est le "nœud" de calcul qui regroupe 8 processeurs double-cœur. Il s'agit d'un serveur Bull Novascale. Les serveurs sont regroupés par 3 dans des armoires : TERA 10 comprend 200 armoires de calcul reliées par 45 commutateurs fournis par la société Quadrics. Ces commutateurs rendent possible la distribution d'une simulation sur tous les processeurs de la machine, ce qui permet donc de disposer, pour une seule simulation, d'une capacité de calcul de plus de 50 000 milliards d'opérations par seconde. Les résultats sont stockés sur 10 000 disques. TERA 10 occupe une surface de 600 m² et consomme environ 1,8 MW.



Plusieurs premières mondiales sont réalisées par cette machine :

- c'est un supercalculateur utilisant *Linux* comme système d'exploitation, marquant l'engagement du CEA et de Bull pour les systèmes "ouverts",
- avec 1 petaoctet (1 million de milliards d'octets) de données directement accessibles, soit 30 fois l'équivalent des informations contenues dans la Bibliothèque Nationale de France, c'est aussi la machine la plus performante pour la gestion des données,
- cette machine, basée sur des "composants sur étagère", est la première à utiliser le dernier né des processeurs *Itanium* d'*Intel*, le "Montecito" double-cœur constitué de 1,7 milliard de transistors.