

Ter@tec, première technopole européenne dédiée à la simulation haute performance.

© www.imaconcepttv

En route pour

Après le pétaflops, c'est l'exaflops (milliard de milliards d'opérations par seconde) qui est visé. Les difficultés à surmonter sont nombreuses.

Léo Gerat,
journaliste scientifique

La barre du pétaflops (million de milliards d'opérations par seconde) a été franchie, en 2008, par le Roadrunner d'IBM, onze ans après le téra-flops (1997 : ASCI Red, Intel). On vise désormais le milliard de milliards d'opérations par seconde. La demande est déjà exprimée par des administrations comme le DOE (*Department of Energy*), aux Etats-Unis, ou certains secteurs scientifiques comme la modélisation climatique.

L'exaflops dans une dizaine d'années ? Autant dire l'infini. Comment gagner ce facteur mille ? Bien peu de réponses, à ce stade, surtout des questions. Un rapport commandé par la Darpa⁽¹⁾ à une équipe dirigée par Peter Kogge (*University of Notre-Dame, Indiana*), intitulé *Exascale Computing Study*, énonce quatre grands "challenges". Le premier est la question de l'énergie. Tout le monde en convient, c'est le problème majeur.

Roadrunner consomme 2,5 mégawatts (MW), ce qui semble déjà une prouesse, puisque Jaguar (Cray) le talonne côté puissance de calcul, mais en consommant 7 MW. S'il fallait pour atteindre l'exaflops multiplier par mille la facture énergétique, il faudrait s'offrir des gigawatts, soit la production d'une grosse centrale. Impensable.

On veut bien à la rigueur imaginer une installation informatique exceptionnelle consommant un dixième de gigawatt. Mais 125 MW, par exemple, pour un exaflops, cela représenterait déjà un progrès en efficacité énergétique d'un facteur 20 par rapport à Roadrunner, ce qui n'est pas rien.

Il va falloir faire de gros progrès dans ce domaine. Cela concerne d'abord les puces qui fourniront les "flops". Or, le coût de développement d'une nouvelle

génération de microprocesseurs est aujourd'hui tel que les fabricants de supercalculateurs n'imaginent plus d'en faire concevoir à leur usage.

« C'est une contrainte avec laquelle il faut se débrouiller, estime William Jalby, responsable de l'équipe Arpa du laboratoire Prism à l'Université Versailles-Saint-Quentin-en-Yvelines. *Seuls des composants "du commerce", produits massivement, offrent un ratio puissance/prix suffisant.* »

Les supercalculateurs actuels emploient des microprocesseurs conçus pour d'autres usages. Il y a peu, c'étaient des puces destinées à des gros ordinateurs, des serveurs. Mais, plus récemment, on a commencé à s'intéresser à des puces plus modestes et économes. Ainsi, le BlueGene d'IBM, numéro un au Top 500 en 2004, utilisait déjà une puce sobre conçue à l'origine pour les systèmes embarqués. Même chose avec le champion de 2008, Roadrunner, toujours chez IBM, qui exploite cette fois une puce conçue pour... une console de jeu, la Playsation de Sony.

« Le monde du calcul scientifique intensif s'habitue à détourner à son profit des puces qui ne lui sont pas destinées, assure Jean-François Lavignon, responsable chez Bull des collaborations R&D. *Nous en avons un bel exemple en France avec ce prototype hybride que Bull réalise pour le Cines (Centre informatique national de l'enseignement supérieur), qui utilise la puce Tesla T10 à 240 cœurs de nVidia, destinée aux cartes graphiques.* »

Voilà pourquoi on dit souvent que les supercalculateurs de la prochaine génération pourraient être réalisés à partir de puces conçues pour des objets grand public nomades. C'est à Santa Clara, en Californie, qu'une équipe de chercheurs du Lawrence Livermore Laboratory et de l'université de Stanford ont trouvé la technologie qui

(1) *Defense Advanced Research Projects Agency*, agence américaine en charge des projets de recherche pour la défense.

L'exaflops



© www.imaconcepttv

Le Campus Ter@tec regroupera 1 000 personnes sur 15 000 m² de laboratoires et de bureaux d'ici à 2011, à Bruyères-le-Châtel.

devrait leur permettre de réaliser un ordinateur exafloppique, capable de simuler le climat terrestre en décomposant l'atmosphère en 20 milliards de cubes.

Xtensa, une puce nomade réalisée par Tensilica, possède 32 cœurs et consomme seulement 0,09 W. D'après les calculs de ces chercheurs, la puissance de calcul par watt obtenue serait plus de quatre fois ce qu'offre la puce du BlueGene d'IBM, et 100 fois supérieure à celle d'une puce Core 2 d'Intel, destinée aux ordinateurs portables.

Le deuxième défi identifié par la Darpa est celui de la mémoire et du stockage. Nourrir des millions de cœurs affamés ne sera pas une petite affaire. Une nouvelle technologie de mémoire serait la bienvenue. Steve Scott, le directeur de la technologie de Cray, est optimiste à ce sujet.

Côté stockage à long terme, la mémoire flash (celle de nos objets nomades), encore chère mais très sobre, devrait jouer un rôle majeur. « On peut s'attendre à l'apparition d'une couche de mémoire flash entre la mémoire centrale et le stockage sur disque », assure Steve Scott. Avis partagé par Steve Pawlowski, le directeur de la technologie d'Intel, un acteur majeur du marché de la mémoire flash.

La troisième gageure identifiée par la Darpa est intitulée : "concurrence et localité". « Il va falloir apprendre à distribuer l'application sur un milliard de cœurs, estime William Jalby. Il n'y aura pas trop de problèmes pour les algorithmes qui travaillent beaucoup sur des données locales. Mais pour les autres... »

« Il y a dix ans, rappelle Franck Cappello, responsable du projet Grand Large à l'Inria Saclay, personne n'aurait prédit que le multicœur serait un pilier du pétaflops. Aujourd'hui, on ne sait pas quels seront les piliers de l'exaflops, et si le multicœur jouera un rôle positif. Il n'est même pas certain qu'il soit si évident à exploiter. »

Du coup, on pense qu'il va falloir littéralement réinventer une nouvelle manière de programmer. On

risque même de voir se répandre des pratiques étranges. « Par exemple, indique William Jalby, il arrivera que l'on préfère recalculer une donnée dont on a un besoin immédiat, plutôt que d'attendre quelle parvienne d'une puce lointaine. »

La gestion de l'énergie elle-même pourrait influencer la programmation. « Si les résultats de deux calculs parallèles sont attendus pour faire autre chose, explique Serge Petiton, responsable de l'équipe MAP au Laboratoire d'informatique fondamentale de Lille, il sera parfois logique de ralentir le plus rapide des deux, pour gagner des watts. »

Le quatrième et dernier pari de l'exascale identifié par la Darpa est la fiabilité. « Une question centrale, estime Jean-François Lavignon, qui nous oblige à réinventer notre façon de concevoir

un ordinateur. » « La complexité de ces machines sera telle, avance Jean-Pierre Panziera, ingénieur chez Silicon Graphics (SGI), que le temps moyen entre deux pannes devrait se mesurer en heures, si ce n'est pire. »

« La panne deviendra un événement normal, estime Franck Cappello. Il faut donc inventer des technologies qui la banalisent, qui permettent de poursuivre le déroulement de chaque logiciel et donc des applications, tandis que l'on répare localement. »

Il va falloir littéralement réinventer une nouvelle manière de programmer



Ter@tec, vue d'architecte.

© www.imaconcepttv