

JC Baratault (NVIDIA) : « NVIDIA a pris la mesure des challenges actuels dans le GPU Computing »

Le **GPU Computing** permet d'exploiter les processeurs des cartes graphiques dans le cadre de calculs classiques. Cette technologie est adaptée aux calculs massivement parallèles, où la puissance exploitable est tout bonnement colossale.

Nous nous sommes entretenus avec **Jean-Christophe Baratault**, EMEA GPU Computing *sales manager* chez NVIDIA.

Cette technologie est dorénavant accessible au travers de bibliothèques standardisées, comme l'OpenCL ou DirectCompute. « *Ce sont des offres bien adaptées aux applications grand public, concède Jean-Christophe Baratault. Nous comprenons parfaitement que certains développeurs préfèrent privilégier une solution standardisée comme l'OpenCL plutôt que CUDA. Toutefois, cette volonté de portabilité entre des architectures différentes rend les optimisations plus difficiles. Certains professionnels préféreront donc des offres plus proches du hardware, comme CUDA, qui permet de tirer la quintessence de nos produits. Nous plaçons CUDA, l'OpenCL et DirectCompute au même plan, tout comme nous travaillons de façon identique dans le domaine de la 3D, avec l'OpenGL, DirectX et Cg.* »

Hier le C, aujourd'hui Fortran... et demain ? Lorsque nous avons lancé le GPU Computing, il a fallu trancher. CUDA, qui se présente sous la forme d'un dérivé du C, semblait un choix pertinent pour attirer un large ensemble de développeurs. Mais nous effectuons un important travail de fond en coopération avec les éditeurs de logiciels et de solutions de programmation. Récemment, des offres Fortran ont été ajoutées. De nouvelles solutions apparaissent : celle d'ActiveEon permet d'adapter Java au marché du cloud computing, de la programmation parallèle et du GPU Computing.



Que pensez-vous de la HighStation XLR8 de Carri Systems ?

Les supercalculateurs NVIDIA Tesla ont suivi les évolutions des cartes mères. Initialement limités à

deux cartes TESLA, ils ont pu passer à quatre, et maintenant à huit avec cette offre de Carri Systems. Nous atteignons toutefois les limites en termes d'encombrement et de bruit. Une autre solution réside dans l'adoption de cartes de calcul au refroidissement passif, ce qui permet de créer des solutions de haute densité, comme le Tesla M1060. Ce sont deux approches différentes, qui ne se destinent pas aux mêmes marchés, ni à la même clientèle. **Pouvez-vous nous en dire plus sur l'architecture Fermi ?** Pas encore ! Nous avons profité de la GPU Technology Conference pour dévoiler cette technologie. Nous présenterons la nouvelle gamme 2010 vers la mi-novembre, lors de la SC09 (Supercomputing Conference) de Portland. Deux nouveautés essentielles sont intégrées à Fermi : le support de la mémoire ECC et la multiplication des cœurs dédiés aux calculs sur les nombres flottants en double précision (dont le nombre passe de 30 à 256). Le support du '64 bits ECC' répond à une forte demande.

Intel et AMD proposent à la fois des CPU et des GPU. Vous sentez-vous en danger ? Nous sommes en danger depuis le premier jour, le marché des cartes graphiques étant particulièrement concurrentiel. Aujourd'hui, nous sommes un des seuls constructeurs indépendants face à ces deux géants. C'est un danger, mais aussi un élément qui a motivé l'ouverture du GPU vers le marché du calcul. Cela n'a pas été aisé : depuis trois ans, nous supportons une charge en terme de coût de production, liée à l'intégration de ces fonctionnalités. Il a aussi fallu convaincre les éditeurs que le GPU Computing présentait des avantages concrets.

Le x86 semble toutefois incontournable, non ? Nous traversons une période charnière, où tout évolue rapidement. Le Tegra souhaite répondre à cette tendance lourde du marché. Reste à savoir comment se placera le x86 sur ces secteurs de gros volume. Cette architecture pourra-t-elle apporter une réponse aux besoins futurs ? Voilà une question intéressante. Au final, les concepteurs de processeurs x86, comme Intel ou AMD, sont face au même challenge que NVIDIA. Une chose est certaine ; quelle que soit la direction que prendra le marché, nous sommes prêts à nous adapter.

[Notre article portant sur la HighStation XLR8 de Carri Systems](#) comportait une petite erreur : en adoptant de nouvelles cartes pourvues de l'architecture Fermi, elle présentera une puissance de calcul en double précision de près de 5 téraflops et non pas de 2,5 téraflops. Le GPU Computing pourrait donc déferler plus vite que prévu au sein du TOP500 des ordinateurs les plus rapides de la planète.