

Autopsie de Tera 10, supercalculateur signé

Bull

Avec l'abandon des essais en réel des bombes nucléaires, la France s'est dotée du plus puissant supercalculateur d'Europe (quatrième au Top 500 mondial), principalement destiné à la simulation pour la sûreté et la fiabilité de l'arme nucléaire.

Bull, unique constructeur européen, a remporté l'appel d'offre. Tera 10 est un puissant **cluster de 602 serveurs Bull NovaScale**. 544 serveurs 'nuds' de calcul disposent chacun de 16 processeurs Intel Itanium 2, répartis dans 270 baies. 56 serveurs sont consacrés aux entrées-sorties et 2 serveurs à l'administration. Le **processeur Intel Itanium 2 'Montecito'** fait son apparition pour la première fois en application. Il n'est pas encore commercialisé par le fondeur ! Sur le calcul, 4352 processeurs dual core fournissent 8704 cœurs. Pour une puissance de calcul de plus de 50 Teraflops (50.000 milliards d'opérations par seconde). Un processeur Montecito réunit 1,7 milliard de transistors. Cette composition technologique fait appel au concept des 'composants sur étagères', des Cots ou '*component off the shelf*'). Une stratégie qui abandonne les processeurs spécifiques et spécialisés ? le CEA était jusqu'à présent équipé de calculateurs Cray ? au profit de composants standards fabriqués en très grande série, sans doute moins performants, mais 100 fois moins chers. Pour répondre aux contraintes imposées par l'appel d'offre du CEA, Bull a adapté l'architecture FAME de ses serveurs NovaScale. Les échanges entre les blocs de processeurs mémoire (QBB) sont assurés via le **chip FSS** (*Fame Scalability Switch*), développé par le constructeur. FSS assure la cohérence globale de la mémoire et des caches, et synchronise l'ensemble des échanges. Chaque serveur comporte plusieurs FSS qui lui permettent d'obtenir, outre des débits importants, une très grande bande passante répondant aux besoins d'entrées-sorties. La **mémoire centrale** est de 30 téra octets distribués, à raison de 48 à 128 Go par nud. La partie **réseau** a été confiée à la société anglo-italienne Quadrics. Ce choix d'une technologie haut débit répond à deux objectifs du CEA : un temps de latence de 3µs (micro seconde) et une capacité d'échange de 650 Go/s, ou la valeur de 160 millions de pages de texte à la seconde. La **bande passante** est de 100 Go/s, soit l'équivalent de 200.000 films en streaming vidéo diffusés en simultané. Indispensable, le **stockage** atteint le Péta octet, ou 1 million de milliards d'octets, directement accessibles, sur 7800 disques. Cette capacité est équivalente à 30 fois la capacité de la *Très Grande Bibliothèque*, ou 250 milliards de pages de texte, ou encore 250.000 films au format DVD. La multiplication des disques n'est pas seulement liée à la capacité de stockage. En exploitant simultanément plusieurs disques, elle permet d'obtenir un débit supérieur à celui d'un seul disque. Sur une simulation où la lecture et l'écriture des données peut atteindre 100 Go par seconde avec Tera 10, le débit d'un disque ne dépasse pas les 40 Mo/s ! L'autonomie de Tera 10 en cas d'accident ou de coupure est de 15 minutes, à peine le temps de sauvegarder ses données par un backup sur bandes. Enfin, le choix du CEA s'est porté sur le logiciel libre. Un choix dicté tout d'abord par l'ouverture, pour partager les développements avec d'autres laboratoires. Mais aussi par la pérennité avec Linux, disponible sur un grand nombre de plateformes et avec une durée de vie qui s'annonce longue. Sans oublier les coûts réduits. Bull s'est engagé lui aussi sur l'open source. Il a modifié le code du noyau de Linux à partir d'une distribution Red Hat, afin de l'adapter à sa technologie HPC. Il a adopté une bibliothèque de communication et le système de fichier global et parallèle Lustre. Et enfin

développé un 'cluster management' pour simplifier la gestion de l'ensemble. Tera 10 pourra traiter quotidiennement entre 10 et 30 téra octets de données? Cette puissance a un prix, 50 millions d'euros. Auxquels il faudra ajouter le coût de 60 ingénieurs pour la maintenance du système, 100 ingénieurs pour les développements et 800 ingénieurs pour son utilisation.