

Introduction à la programmation des clusters

Introduction

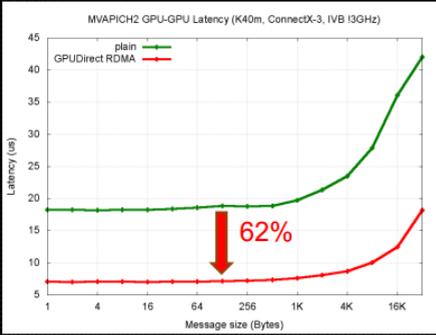
Le GPU Computing sur cluster consiste en réalité à une couche supplémentaire par rapport au calcul sur un cluster classique. Dans ce dernier cas, vous vous baserez principalement sur la bibliothèque MPI pour communiquer entre les nœuds du cluster. Il y aura donc un niveau de communication supplémentaire, du GPU vers le CPU, puis entre les CPU puis à nouveau vers le GPU. Cette quantité d'allers-retours va être le goulot d'étranglement de votre application et il sera donc indispensable de minimiser ces communications et de les optimiser. Une manière de les optimiser est de placer correctement les processus sur les nœuds afin que les communications les plus fréquentes se produisent entre cores qui sont proches.

Pour éviter l'étape GPU vers CPU, NVIDIA a mis en place une option dans la version XX (je crois, avec GPU Direct) de CUDA permettant de transférer directement par MPI des données entre GPU. Dans ce cas, il n'y a pas de copie intermédiaire sur CPU, mais toutes les implémentations d'MPI ne supportent pas ce mode. Les fournisseurs de clusters ne le proposent généralement pas, il faudra faire un compromis entre l'implémentation préconisée par votre prestataire et une implémentation Open Source comme OpenMPI qui n'est peut-être pas autant optimisée.

GPU Direct

Permet le transfert directe GPU-réseau

● Reduced inter-node latency



● Better MPI Application Scaling

