

Point sur la recherche en calcul haute performance

Thierry Priol

IRISA/INRIA

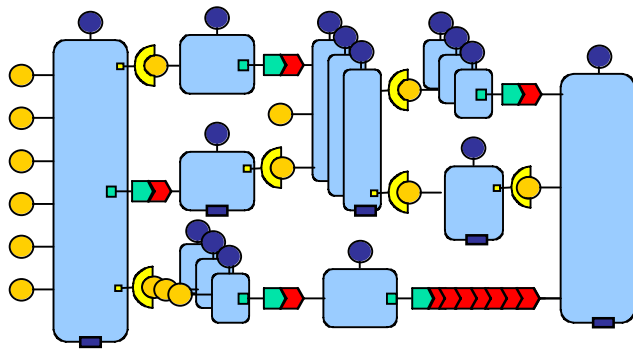
E-Mail : Thierry.Priol@irisa.fr

Le calcul haute-performance s'est toujours appuyé sur des architectures vectorielles et/ou parallèles homogènes. Avec les évolutions des technologies réseaux à très haut débit, dûes aux succès de l'Internet, il est désormais possible de construire de nouvelles infrastructures de calcul rassemblant plusieurs calculateurs parallèles répartis géographiquement et interconnectés par des réseaux à très haut-débit (de 10 à 40 Gbit/s aujourd'hui). Ces infrastructures cumulent les capacités de calcul et de stockage des calculateurs parallèles mis ainsi en réseau. Ces infrastructures ont été popularisées par le concept de grilles informatiques, idée qui signifie une globalisation et une dématérialisation des ressources informatiques. Un utilisateur peut ainsi utiliser des ressources disponibles sur le réseau sans savoir d'où elles proviennent. La réalisation de grilles informatiques se heurte à de nombreuses difficultés d'ordre scientifique et technique. Il s'agit de faire communiquer et coopérer, à grande échelle, des matériels distants et hétérogènes par leurs modes de fonctionnement comme par leurs performances, de créer les logiciels permettant de gérer et de prendre en compte efficacement la distribution des équipements, de mettre au point des outils de programmation adaptés au caractère diffus et à grande échelle de l'exécution des tâches confiées à la grille, etc.

Au sein du projet PARIS de l'IRISA, nous nous intéressons notamment aux modèles de programmation qui permettent de programmer les grilles informatiques en utilisant un niveau d'abstraction suffisant pour que le programmeur soit capable d'exploiter facilement ce type d'infrastructure pour des applications du calcul scientifique (comme les applications de couplage de codes qui mettent en jeu plusieurs applications simulant des phénomènes physiques qui interagissent entre eux).

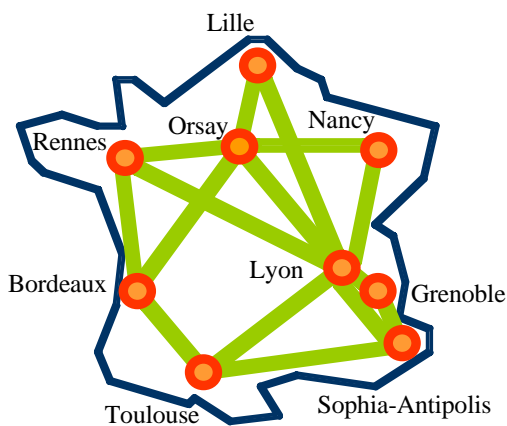
QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.

Les techniques classiques pour coupler plusieurs codes de calcul sont fondées sur l'utilisation de bibliothèques de communication telles que MPI ou PVM. Elles sont soit exposées directement au programmeur ou bien intégrées au sein de coupleurs spécialisés (Calcium, MpCCI, ...). Ces approches ont cependant l'inconvénient d'être fondées sur des techniques issues du calcul parallèle. Pour les grilles informatiques, le déploiement d'applications de couplage de codes pose de nombreux problèmes : mise en œuvre des bibliothèques de communication pour la grille capable de prendre en compte le caractère hétérogène des ressources, déploiement des codes, etc... Les recherches, que nous avons entreprises depuis plusieurs années au sein du projet PARIS, sont plutôt fondées sur une approche par composants logiciels qui intuitivement semble appropriée pour exprimer le couplage de plusieurs codes de calcul. Chaque code est encapsulé au sein d'un composant logiciel et l'application de couplage est alors un assemblage de composants logiciels. Cependant, les modèles de composants logiciels ne sont pas adaptés au contexte du calcul haute-performance et ceci pour plusieurs raisons.



Les communications entre composants logiciels ne sont pas très efficaces car elles sont souvent fondées sur le protocole de communication TCP/IP qui ne sait pas tirer parti de réseaux plus rapides au sein des calculateurs parallèles. Les modèles de composant logiciels ne sont pas adaptés à l'encapsulation de codes parallèles. La communication entre deux composants logiciels qui encapsulent des codes parallèles ne peut exploiter les débits offerts

par les réseaux à très haut-débit actuellement disponible dans l'Internet. Nos recherches ont montré que tous ces obstacles peuvent être surmontés et qu'il est possible d'utiliser une approche par composant logiciels pour le couplage de codes dans un contexte calcul haute-performance.



Des expériences ont été réalisées sur une infrastructure de grilles informatiques qui est actuellement en cours de montage : il s'agit du projet Grid'5000 soutenu par le Ministère à l'Enseignement Supérieur et à la Recherche dans le cadre de l'ACI GRID avec le concours de l'INRIA, du CNRS, de Renater, des Universités et des financements de conseils généraux et régionaux. Grid'5000 a pour objectif la construction d'un instrument de recherche pour les grilles informatiques rassemblant des grappes de PC sur neuf sites en France (dont l'IRISA à Rennes) avec environ 5000 processeurs. Les neuf sites sont interconnectés par le réseau Renater qui a mis à disposition une connectivité réseaux (10 Gbit/s)¹. En février 2006, 1800 processeurs sont d'ores et déjà

installés. À l'horizon 2007, l'instrument Grid'5000 possèdera environ 3500 processeurs et devrait atteindre en 2008 environ 5000 processeurs. Le nœud Rennais de Grid'5000 possède 530 processeurs offrant ainsi une puissance crête de l'ordre de 2 Teraflops. Cette grappe possède également des réseaux rapides (Infiniband, Myrinet 2000, Myrinet 10G).



¹ Pour des raisons de sécurité, l'instrument Grid'5000 est physiquement isolé par Renater afin d'empêcher toute intrusion externe.