

# Recouvrement des calculs et des communications MPI

Alexandre DENIS  
Alexandre.Denis@inria.fr

Octobre 2013

Les supercalculateurs sont un assemblage de noeuds de calcul reliés par un réseau rapide. Au niveau applicatif, le réseau est utilisé au travers de l'interface MPI. Pour amortir le temps nécessaire aux communications, il est d'usage courant de recouvrir le temps de communication par des calculs en utilisant les primitives non-bloquantes de MPI. Cependant, la norme MPI ne garantit pas que quand une communication est lancée par une primitive non-bloquante, elle va effectivement progresser en tâche de fond.

L'objectif du sujet d'Études et de Recherche proposé est de :

- Faire une synthèse de la problématique en utilisant les références ci-dessous ;
- Expliquer la solution proposée dans l'article [1].
- Évaluer le comportement de différentes implémentations MPI courantes par rapport au recouvrement des calculs et des communications.

## Références bibliographiques

1. Mohammad Rashti et Ahmad Afsahi. *Improving communication progress and overlap in MPI rendezvous protocol over rdma-enabled interconnects*. In High Performance Computing Systems and Applications, 2008. HPCS 2008. IEEE.
2. Mohammad Rashti et Ahmad Afsahi. *Assessing the ability of computation/communication overlap and communication progress in modern interconnects*. In High-Performance Interconnects, 2007. HOTI 2007. IEEE.
3. François Trahay and Alexandre Denis. *A scalable and generic task scheduling system for communication libraries*. In Proceedings of the IEEE International Conference on Cluster Computing, New Orleans, LA, September 2009. IEEE Computer Society Press.

4. Torsten Hoefler and Andrew Lumsdaine. *Message Progression in Parallel Computing - To Thread or not to Thread?*. IEEE Cluster 2008 Tsukuba, Japan, 2008.