

Structure de DAG pour le calcul parallèle en Python

ENCADRANT : Sébastien Limet

Sujet

Dans de nombreux domaines scientifiques, la structure de DAG (*Directed Acyclique graph*) est utilisée pour faire des simulations numériques. Les DAG sont utilisés pour représenter des systèmes sanguins ou des réseaux hydrauliques. Dans ce type de simulations, des données sont associées aux arcs, et d'autres aux nœuds. De même, une simulation dirige l'évolution des données sur les arcs (l'écoulement du sang dans les artères par exemple) et une autre simulation permet de représenter ce qui se passe au niveau des jonctions (par exemple comment le sang se répartit dans les artères).

Pour permettre de simplifier la programmation de telles simulations, Hélène Coullon dans le cadre de sa thèse a développé une structure de données permettant

- de distribuer une telle structure de DAG et les données qui y sont associées sur une machine parallèle,
- de définir les traitements à appliquer sur les arcs et sur les nœuds,
- de lancer le calcul parallèle

La structure de données est optimisée pour permettre d'ordonner les calculs de telle sorte que l'on puisse optimiser le calcul grâce au recouvrement communications/calculs.

L'objectif de ce TER est d'implémenter un prototype en Python de la structure de DAG définie dans la thèse d'Hélène Coullon. Il s'agit d'utiliser des bibliothèques Python qui permettent d'optimiser les étapes de calculs et de prendre en charge la parallélisation du code. Ainsi, ce prototype doit permettre de définir une simulation sur un DAG, de charger et répartir le DAG sur une machine parallèle et enfin de lancer la simulation.

Livrables attendus

- L'implémentation de l'environnement permettant de définir une simulation sur un DAG
- L'implémentation d'une simulation grâce à cet environnement
- Une étude de la performance de l'implémentation tant en termes de temps de calcul que du passage à l'échelle.