

paraloop : une boucle parallèle intégrée à une chaîne de traitements

Emmanuel COURCELLE et Jérôme GOUZY
L.I.P.M. (UMR CNRS-INRA 2594/441) – B.P.52627
31326 CASTANET TOLOSAN Cedex – FRANCE

Mots clefs

Parallélisation à gros grain, calcul scientifique

Résumé

Les architectures parallèles se sont banalisées depuis quelques années, au sein même des laboratoires, tout d'abord sous la forme de machines multiprocesseurs, puis plus récemment sous la forme de grappes de calcul. Elles sont utilisées pour exécuter des calculs lourds via une chaîne de traitements constituée d'un script enchaînant lui-même plusieurs scripts ou programmes binaires.

Certains traitements, parmi les plus lourds, se résument à *i*) l'ouverture d'une ressource (base de données, fichier, ...) constituée de nombreux enregistrements, *ii*) l'exécution d'un traitement sur chacun de ces enregistrement, *iii*) l'écriture du résultat dans un fichier de sortie.

Ce type de traitements se prête bien à une parallélisation à gros grains, dans la mesure où il n'y a pas de dépendance de données entre les enregistrements. Toutefois, il n'est pas rare de devoir écrire plusieurs fois le même script, en raison de la diversité des architectures parallèles accessibles: machines à mémoire partagée ou grappes de calculateurs, installations pourvues ou non d'un système de queue, systèmes de queue configurés différemment, etc.

Paraloop, écrit en perl orienté objets, a pour objectif de masquer ces différences, en encapsulant dans un objet le code correspondant à l'architecture de la machine utilisée. Ainsi, il est possible d'exécuter la même chaîne de traitements sur différentes architectures sans changer une ligne à son programme, simplement en modifiant quelques paramètres dans un fichier de configuration.

Issu du monde de la Bioinformatique, *paraloop* se révèle très utile pour tous les projets scientifiques dans lesquels un parallélisme à gros grain est envisageable.

Paraloop, couvert par la licence Cecill v2, peut être téléchargé à l'adresse <http://lipm-bioinfo.toulouse.inra.fr/paraloop>