



Simulation et outils de debug pour réseaux de processus

Matthieu Moy – Maître de conférences Université Lyon1/LIP
Christophe Alias – Chargé de recherche INRIA/LIP

Mots clés Simulation, Parallélisme, SystemC, Modèle flot de données, Modèle Polyédrique

Lieu Laboratoire de l'Informatique du Parallélisme (LIP)
École Normale Supérieure de Lyon

Contexte

Depuis le début des années 2000, la limite de miniaturisation des transistors force à multiplier les unités de calcul (processeurs, processeurs spécialisés) des superordinateurs pour améliorer les performances [1], ce qui augmente la consommation énergétique et donc le coût du calcul dans des proportions qui ne sont plus tenables. De nouveaux paradigmes de calcul basse consommation doivent être inventés. Une solution est de cabler directement les algorithmes principaux du calcul sur des circuits reconfigurables FPGA (Field Programmable Gate Array) [2]. Pour exécuter une application sur des FPGA, de nombreux verrous doivent être levés, dont la traduction automatique d'un algorithme de calcul en un circuit efficace (HLS, High-level synthesis).

La traduction d'un programme en circuit se fait en deux étapes. Tout d'abord, le *front-end* produit une représentation intermédiaire adaptée à la synthèse de circuit. Dans les outils CASH, ce formalisme s'appelle DPN [3], pour « Data-aware Process Network », et représente un réseau de processus qui capture le parallélisme de l'application et les communications entre unités parallèles. Ensuite, le *back-end* traduit chaque composant du réseau de processus en matériel en assurant une bonne réutilisation des ressources. Au final, le circuit produit peut être vu comme un gigantesque réseau de processus pipeliné, acceptant des entrées et produisant des sorties à intervalles réguliers.

L'équipe CASH nouvellement créée travaille sur des approches innovantes d'extraction du parallélisme vers une représentation intermédiaire. La génération de code finale vers le FPGA est déléguée à un outil de synthèse externe : nous utilisons donc des compilateurs dits « source-to-source », qui lisent du code C séquentiel, en extraient le parallélisme sous forme d'un réseau de processus, puis génèrent du code qui explicite le parallélisme.

Pour le développement et le debug des réseaux de processus utilisés comme représentation intermédiaire (DPN), il est nécessaire de pouvoir simuler ces réseaux de processus sans passer par l'étape de synthèse (très longue) et l'exécution sur FPGA (qui a des capacités de debug très limitées). Nous utilisons pour l'instant un simulateur basé sur les threads POSIX, en utilisant un thread par processus du réseau. Cette solution est opérationnelle, mais peu pratique pour le debug, et assez lente du fait de la fréquence des changements de contextes.

Une alternative serait d'utiliser un simulateur dédié à la simulation de matériel, comme SystemC. SystemC est une bibliothèque C++ qui contient un noyau de simulation basé sur un ordonnanceur coopératif. Il est fourni avec un grand nombre de primitives de synchronisation et de communication. C'est le standard industriel pour la simulation rapide de matériel.

Objectifs du TER

L'objectif général du TER est d'étudier la faisabilité et les bénéfices potentiels d'un simulateur basé sur SystemC.

Plus précisément, les objectifs du TER sont :

- Comprendre le contexte, le fonctionnement général d'un synthétiseur de haut niveau (HLS) ;
- Déterminer un sous-ensemble du formalisme DPN qui soit suffisant pour représenter des programmes non-triviaux, mais suffisamment simple pour permettre l'implémentation d'un simulateur dans le cadre du projet ;
- Implémenter un simulateur pour ce sous-ensemble basé sur SystemC.

En fonction des premiers résultats, nous pourrions également considérer :

- Des outils de visualisation permettant d'examiner des traces d'exécution du simulateur ;
- La parallélisation du simulateur basé sur SystemC. En effet, SystemC simule le parallélisme, mais le fait de manière séquentielle. La parallélisation de SystemC est un sujet de recherche actif et la forme particulière des DPN devrait permettre de bonnes performances en parallélisme.

Encadrement

Ce projet sera co-encadré par Christophe Alias (CR1 Inria, ENS-Lyon) et Matthieu Moy (MCF HDR UCBL).

Christophe Alias (<http://perso.ens-lyon.fr/christophe.alias/>) s'intéresse à la synthèse de circuit haut niveau dans le modèle polyédrique depuis plus de 8 ans. Il a co-encadré deux thèses (Alexandru Plesco avec Alain Darte et Tanguy Risset, et Guillaume Iooss avec Sanjay Rajopadhye). Dans le même temps, il a écrit un compilateur de réseaux de processus, transféré dans la startup Xtremlogic qu'il a co-fondé en 2014 avec Alexandru Plesco. Actuellement Christophe Alias est en concours scientifique à 20% dans XtremLogic et n'a plus de charge d'encadrement depuis la soutenance de Guillaume Iooss en juillet 2016.

Matthieu Moy (<https://matthieu-moy.fr>) travaille sur la simulation en SystemC depuis une quinzaine d'années (en partenariat avec STMicroelectronics et en particulier dans le cadre de la HLS), et a déjà encadré plusieurs thèses et post-doctorants sur le sujet. Plus récemment, il s'est intéressé aux calculs de pire temps d'exécution de logiciel et de pire temps de traversée de réseaux sur puces dans le cadre de systèmes temps-réel critiques. Il est titulaire de l'habilitation à diriger des recherches depuis 2014. Anciennement responsable de l'équipe Synchronique du laboratoire Verimag, il a intégré le LIP en septembre 2017. Il co-encadre aujourd'hui 4 thèses.

Compétences souhaitées

Notions solides en parallélisme, notions en architecture des ordinateurs. Maîtrise du langage C, bases solides en C++.

Candidatures

Envoyez vos candidatures par email à : Christophe.Alias@ens-lyon.fr et Matthieu.Moy@univ-lyon1.fr.

Références

- [1] Nor Zaidi Haron and Said Hamdioui. Why is cmos scaling coming to an end? In *Design and Test Workshop, 2008. IDT 2008. 3rd International*, pages 98–103. IEEE, 2008.
- [2] Altera Corporation. Altera FPGAs achieve compelling performance-per-watt in cloud data center acceleration using CNN algorithms. <http://www.prnewswire.com/news-releases/altera-fpgas-achieve-compelling-performance-per-watt-in-cloud-data-center-acceleration-using-cnn.html>, 2015.
- [3] Christophe Alias and Alexandru Plesco. Data-aware Process Networks. Research Report RR-8735, Inria - Research Centre Grenoble – Rhône-Alpes, June 2015.