

Sujet de thèse : Exploration optimisée pour la conception automatisée des futurs processeurs

Encadrants au CEA-LIST: Lilia Zaourar, Marc Duranton

Encadrant universitaire : Alix Munier Kordon

1 Description du sujet

La conception des nouvelles générations de processeurs dotés de multiples coeurs de calcul et d'interfaces rapides, est un challenge. Dans ce contexte, le développement d'outils d'aide à la décision pour la conception de ces processeurs est indispensable et présente de nombreux défis, ce qui en fait un champ de recherche très actif [1]. L'objectif de cette thèse est de spécifier, implémenter et expérimenter des algorithmes d'exploration automatique pour la conception de processeurs hétérogènes [2]. Ces algorithmes seront intégrés au framework A-DECA (Automated Design space Exploration for Computing Architectures) développée au Laboratoire Environnement de Conception et Architecture (LECA) du CEA-LIST. Une première solution à base d'algorithmes génétiques multi-objectifs existe déjà. Elle permet d'obtenir un ensemble non dominés de configurations pour des exemples d'architectures multi-coeurs homogènes.

Le point départ de cette étude est l'analyse des résultats de simulations de l'exécution d'un code donné sur une architecture matérielle définie par à un ensemble de paramètres: nombre et type de coeurs, hiérarchie mémoire, dimensionnement des caches, réseaux d'interconnexion etc...). Ces simulations fournissent un ensemble de mesures de performance comme le temps d'exécution de l'application sur l'architecture choisie, la surface de l'architecture..etc.. Les problèmes d'ordonnancement et de placement traités par ces simulateurs sont difficiles [3, 4]. Les algorithmes utilisés intègrent un grand nombre de contraintes et leurs résultats est difficilement prévisible avec finesse. Le premier travail va ainsi consister à étudier les résultats fournis par ces simulateurs pour proposer des modèles de prédiction, et évaluer leur fiabilité. Il s'agit ici d'effectuer une étude bibliographique sur différentes techniques de régression (modèles de régression linéaires, régressions logistiques [5], ou des techniques de machine learning [6]) pour déterminer et tester expérimentalement les méthodes retenues pour à la fois réduire les paramètres significatifs de l'architecture matérielle et prédire les résultats des simulations non effectuées. Pour cela, le candidat sera amené à prendre en main l'environnement et à s'approprier l'infrastructure existante avec des exemples d'applications.

Dans un second temps, le candidat pourra s'appuyer sur des travaux existant dans l'équipe afin d'étudier l'intégration d'algorithmes d'exploration de l'espace de conception dans A-DECA pour réaliser une infrastructure d'exploration. Il s'agit ici de développer et tester un ou plusieurs algorithmes pour calculer des configurations architecturales. Les algorithmes s'appuieront sur les prédicteurs développés précédemment pour évaluer les solutions proposées.

Enfin, le candidat mettra en oeuvre ses développements en réalisant une exploration de l'espace de conception d'une architecture de calcul à l'aide d'une ou plusieurs applications plus complexes provenant par exemple de l'intelligence artificielle ou du HPC afin de déduire des bonnes configurations architecturales.

2 Contexte de l'étude

La thèse sera préparée au sein du Laboratoire Environnement de Conception et Architecture (LECA) qui est impliqué dans plusieurs projets de conception ou d'aide à la conception de circuits électroniques complexes. Le LECA est ainsi fortement impliqué dans le projet de processeur européen EPI (European Processor Initiative). L'étudiant sera hébergé au CEA LIST, au département DSCIN. Paris-Saclay Campus - Nano-INNOV, Bât. 862-PC172 F-91191 Gif-sur-Yvette Cedex.

L'inscription universitaire s'effectuera à Sorbonne université, et le co-encadrement scientifique sera réalisée en lien avec l'équipe ALSOC du LIP6, 4 place Jussieu, 75252 Paris cedex 05. Une présence est donc nécessaire pour participer à la vie de l'école doctorale et de l'équipe de recherche (séminaires, formations, etc.) ainsi que les échanges scientifiques avec la directrice de thèse.

Contacts: lilia.zaourar@cea.fr, alix.munier@lip6.fr

3 Profil souhaité

Le candidat recherché est en dernière année de master recherche ou diplôme ingénieur (bac+5). Des connaissances solides en optimisation combinatoire algorithmique, recherche opérationnelle, machine learning, langages C/C++ et Python sont requises. Des connaissances en architectures de calcul, processeurs multi-coeurs seront appréciées.

Exigeant et investi, vous avez à cœur de proposer des solutions innovantes et de travailler dans un milieu à la pointe de la technologie qui vous permettra de répondre aux enjeux de demain. Le candidat devra être doté d'un bon relationnel et posséder la capacité de travailler en équipe et en autonomie.

Documents: CV détaillé, lettre de motivation, notes et rang sur les 3 dernières années, le nom d'un référent qui peut vous recommander.

Références

- [1] M. Duranton, K. D. Bosschere, B. Coppens, C. Gamrat, T. Hoberg, H. Munk, C. Roderick, T. Vardanega, and O. Zendra, "The hipec vision for advanced computing in horizon 2020," 2021.
- [2] M. Duranton, J. Hoogerbrugge, G. Al-kadi, S. Guntur, and A. Terechko, "Rapid technology-aware design space exploration for embedded heterogeneous multiprocessors," in *Processor and System-on-Chip Simulation*, pp. 259–275, Springer, 2010.
- [3] M. A. Aba, L. Zaourar, and A. Munier, "Efficient algorithm for scheduling parallel applications on hybrid multicore machines with communications delays and energy constraint," *Concurr. Comput. Pract. Exp.*, vol. 32, no. 15, 2020.
- [4] L. Zaourar, M. Ait Aba, D. Briand, and J.-M. Philippe, "Task management on fully heterogeneous micro-server system: modeling and resolution strategies," *Concurrency and Computation: Practice and Experience*, vol. 30, no. 23, p. e4798, 2018.
- [5] K. Train, *Discrete Choice Methods with Simulation*. Cambridge University Press, 2009.
- [6] G. Iomazzo, C. D'Ambrosio, A. Frangioni, and L. Liberti, "A learning-based mathematical programming formulation for the automatic configuration of optimization solvers," in *Machine Learning, Optimization, and Data Science - 6th International Conference, LOD 2020, Siena, Italy, July 19-23, 2020, Revised Selected Papers, Part I*, pp. 700–712, 2020.