Multi-Objective Optimization of Speculative and Anticipative Multi-Core Architectures

Multi-Objective Optimization of Speculative and Anticipative Multi-Core Architectures

Claudiu Buduleci, Árpád Gellért, Adrian Florea, Radu Chiș, Remus Brad

University "Lucian Blaga" of Sibiu, Computer Science and Electrical Engineering Department, E. Cioran 4, 550025 Sibiu, Romania, E-mail: claudiu.buduleci@ulbsibiu.ro

ABSTRACT

This paper explores niche multi-core research opportunities. We are proposing speculative and anticipative micro-architectural techniques for a state-of-the-art multi-core system. Further, we propose different multi-objective optimization methods for improving the automatic design space exploration process in multi-core architectures.

KEYWORDS: Multi-Core; Multi-Objective Optimization; Value Prediction; Dynamic Instruction Reuse; Sniper.

1 Introduction

Multi-Objective Optimization (MOO) is an open problem with active involvement from academic and industrial worlds, at an international level. Nowadays, the complexity of problems is increasing, and the time needed for a thoroughly analysis of problems is decreasing. Due to this phenomenon an exhaustive analysis of problems is not possible in most cases (e.g. optimizing micro-processors, electronic control units, cars, etc.). As a consequence, a robust problem-solving technologies and methods are needed. Dealing with multiple objectives (performance, area of integration, energy consumption and thermal dissipation) and a huge design space, leads to the necessity of automatic design space exploration process (ADSE).

This research proposal has the following two goals:

- To develop original effective speculative and anticipative methods/techniques in a state-of-the-art multi-core system;
- To improve the Automatic Design Space Exploration (ADSE) process. New Effective Pareto Multi-Objective Optimization Methods:
    - Use Response Surface Modeling (RSM) techniques;
    - Develop a domain knowledge for multi-core systems using an adequate knowledge representation;
    - Multi-objective optimization through novel pareto-fuzzy genetic algorithms.

2 Methodology

2.1 Implementation of Advanced Micro-Architectural Techniques

A first step will be to implement the proposed multi-value dynamic instruction reuse (DIR) and the value prediction (VP) schemes in a multi-core environment. The fundamental difference between VP and DIR is that the first is a speculative technique and the second is non-speculative (anticipative). This means that the DIR technique eliminates the speculative execution of instructions. Basically, the recovery time payed in case of a wrong prediction does not exist in DIR. This puts DIR in advantage over VP. Both architectural techniques exploit the value locality in a history context of *n* values. Therefore, a deep comparison of DIR vs. VP techniques in fair conditions (to exploit the same value locality degree), in a multi-core system will be very interesting. Their implementation shall be done selectively (on high latency instructions, with focus on critical path, etc.). The selectiveness is needed to consume less energy, because nowadays performance/watt is more important than performance itself. A big challenge will be to implement correctly those techniques, in order to not corrupt the shared variables coherency which are predicted by different execution threads.

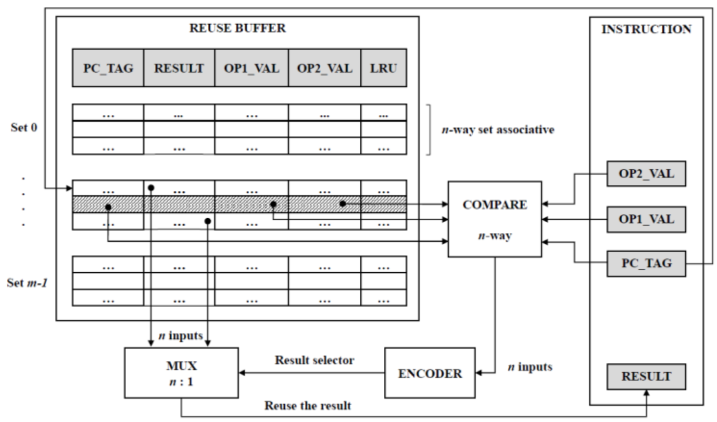

Figure 1 - Multi-Value Dynamic Instruction Reuse

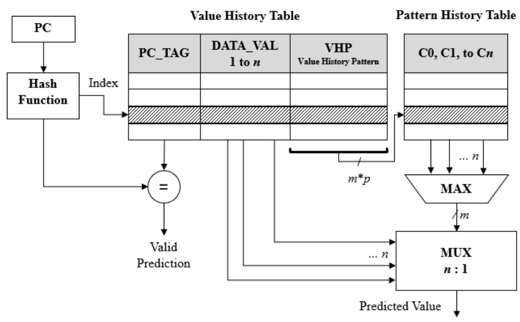

Figure 2 - Value Prediction Based on Multiple Values

An idea is to enhance the reuse buffer (RB) from S_v DIR scheme proposed in [1]. The enhancement consists in proposing to use an *n*-way set associative RB (multi-value). It will store *n* results for one instruction. The proposed scheme is visible in Figure 1. In this way, it can exploit the same *n*-value locality as VP through contextual and computational predictors.

Figure 2 shows a scheme for a generic two-level adaptive value predictor. This scheme is a generalized version of the one proposed in [2]. We are also considering implementing the perceptron based predictors proposed in [3], selective predictor from [4] and the schemes used in the 1st Championship Value Prediction (CVP-1)[1]. Recent research on this topic has been done in [5] and [6].

---

[1] ISCA, "1st Championship Value Prediction." https://www.microarch.org/cvp1/cvp1/ (accessed 18.06.2020).

2.2 Simulation

A state-of-the-art simulator (e.g. Sniper, GEM5) will be chosen and enhanced with the proposed techniques. A devoted simulation methodology will be developed for the comparison of the enhanced micro-architectural techniques, in fair conditions. For interest will be the following metrics: instruction per cycles (IPC), power consumption (total and average), energy consumption (total) and die temperature (average and maximum).

2.3 Optimization

Further elaboration on the simulation results will be done in this phase. The proposed multi-core architectures will be optimized using existing state-of-the-art multi-objective optimization techniques, integrated in an automatic design space exploration process using FADSE [7]. The optimization will be enhanced using response surface modeling techniques, through expert domain knowledge representation and with novel multi-objective genetic algorithms (located in Pareto-Fuzzy paradigm). The optimization performance will be measured using the following metrics: Hypervolume, Two Set Difference Hypervolume (TSDH) and Coverage of Two Sets.

Response Surface Modeling (RSM) represents a collection of statistical and mathematical techniques used to approximate a system. Based on this observation, a response surface (approximation model) of the system is empirically created. It aims to speed up the ADSE process. The computational constraint of the simulation, in terms of time and computational resources, can be significantly reduced. Instead of running an actual simulation we can create approximation functions (response surfaces – RS) and predict the output of the simulator. This approach will allow for sure a deeper exploration concerning the whole design space. Of course, the tradeoff will be the simulation accuracy. The following approaches are considered: classical RS designs (mathematical interpolation based) and Genetically Programmed RS (knowledge discovery).

The optimization algorithm is not specialized in solving a specific problem. They are specified in a generic way. To help the algorithm towards a faster convergence and better quality of solution, domain specific (expert) knowledge can be used. The knowledge can be expressed, e.g. using fuzzy rules expressed in fuzzy logic. The idea is to create a domain-knowledge for a multi-core system. The challenges will be the following: How can a domain knowledge for multi-core system be developed? What are the rules? Are the rules contradictory? How to justify the rules? How to update the mutation and crossover operators? How many rules shall we use? Which rules are more important?

The latest idea is to develop new innovative genetic algorithms, located in the Pareto-Fuzzy paradigm. In specialty literature, several approaches exist, in which an individual from a Pareto crowd is assigned a degree of belongingness to that crowd. In consequence, there are multi-objective genetic algorithms that lay in this Pareto-Fuzzy paradigm.

2.4 Validation

Finally, a validation against the state-of-the-art shall be done. It will be achieved through an ADSE process of the enhanced multi-core architecture, targeting four objectives: performance, area of integration, temperature dissipation and energy consumption. The internally developed FADSE tool will be used, and for thermal estimation we plan to use our methodology developed in [8].

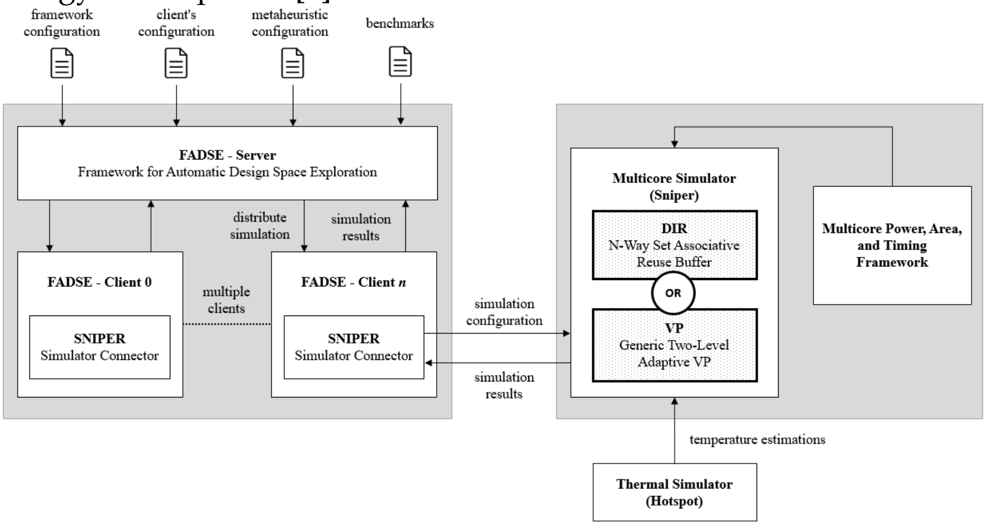

Figure 3 - Methodology overview of the ADSE process

3 Conclusions

In our research group we have expertise in various anticipative and predictive microarchitectural techniques, DSE and thermal evaluation. The highest challenge will be to combine all those ideas originated in mono-core to a multi-core system, considering also a correct execution of the program (e.g. shared data consistency).

4 References

[1] A. SODANI, "Dynamic Instruction Reuse," University of Wisconsin - Madison, 2000.
[2] K. WANG and N. FRANKLIN, "Highly Accurate Data Value Precision Using Hybrid Predictors," *Proc. 30th Int. Symp. Microarchitecture*, 1997.
[3] A. THOMAS and D. KAELI, "Value Prediction with Perceptrons," vol. The 2nd Value-Prediction and Value-Based Optimization Workshop (VPW2) in conjunction with the 11th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 11), Oct. 2004.
[4] A. GELLÉRT and L. VINȚAN, "A multicore architecture with selective load value prediction," *Proc. Romanian Acad. Ser. Math. Phys. Tech. Sci. Inf. Sci.*, vol. 19, no. 4, 2018.
[5] L. OROSA, R. AZEVEDO, and O. MUTLU, "AVPP: Address-first Value-next Predictor with Value Prefetching for Improving the Efficiency of Load Value Prediction," *ACM Trans. Archit. Code Optim.*, vol. 15, no. 4, pp. 1–30, Jan. 2019, doi: 10.1145/3239567.
[6] A. PERAIS and A. SEZNEC, "EOLE: Combining Static and Dynamic Scheduling Through Value Prediction to Reduce Complexity and Increase Performance," *ACM Trans. Comput. Syst.*, vol. 34, no. 2, pp. 1–33, May 2016, doi: 10.1145/2870632.
[7] H. CALBOREAN, "Multi-Objective Optimization of Advanced Computer Architectures using Domain-Knowledge," "Lucian Blaga" University of Sibiu, Faculty of Engineering, Sibiu, 2011.
[8] A. FLOREA, C. BUDULECI, R. CHIȘ, A. GELLÉRT, and L. VINȚAN, "Enhancing the Sniper simulator with thermal measurement," in *2014 18th International Conference on System Theory, Control and Computing (ICSTCC)*, Sinaia, Oct. 2014, pp. 31–36, doi: 10.1109/ICSTCC.2014.6982386.