A NEW LOW POWER EXPLORATION MECHANISM BASED ON DESIGN OF EXPERIMENTS (DOE) AND TWO-LEVEL HIERARCHIES

Cordeiro, F.R. and Silva-Filho, A.G.

Federal University of Pernambuco (UFPE), Informatics Center (CIn)

ABSTRACT

Tuning cache architectures in platforms for embedded applications can dramatically reduce energy consumption. This paper presents an optimization mechanism based on Design of Experiments (DoE) for adjusting two-level cache memory hierarchy in order to reduce energy consumption of embedded applications. DoE is a technique used to plan experiments and in this work it was adapted to architecture exploration problem. Preliminary results for 6 applications from Mibench benchmark suite show an average reduction of about 6% in the energy consumption for data caches and it has revealed simpler with low computational cost, when compared to existing heuristics.

1. INTRODUCTION

The memory subsystem has been demonstrated to be the energy bottleneck: several researchers [1][2] have demonstrated that the memory subsystem now accounts for 50-70% to the total power budget of the system. Thus, many efforts have been made in order to develop optimization mechanisms to reduce energy consumption without degrade the performance. Previous researches have observed that adjustment of the cache parameters to a specific application can save about 60% of energy consumption in the memory system [3].

Memory hierarchies with two cache levels have large exploration space, being able to reach some hundreds of different configurations. Exploration mechanisms based on heuristics [4], can be applied efficiently due its simplicity and sometimes its efficacy. Intelligent mechanisms based on genetic algorithms [5], although it provide a higher number of simulations, it can offers decision support for the hardware designer with good candidates through the Pareto-Front. In the majority of the cases, the strategy used to define and formalize an optimization mechanism is based on simulations and average values. In this work, a strategy based on planning of experiments that allows studying the effect of each analyzed parameter, as well as the effects of interactions between then on a given constraint is proposed.

In this paper, an optimization mechanism for adjusting two-level data caches hierarchies, called TECH-DoE (Two-level cache Exploration Mechanism based on Design of Experiments), is proposed.

2. RELATED WORK

Nowadays, many efforts have been made in order to reduce the consumption of energy by adjusting cache parameters to the needs of a particular application. However, contributions for tuning two-level cache hierarchies, with separated instruction and data caches for both levels, do not consider a statistical model to analyze the behavior of the application.

Gordon-Ross et al. [6] has extended the Zhang heuristics, intended for one-level caches, and proposed the TCaT heuristic, intended for two-level hierarchy. The use of TCaT heuristic allows energy savings of 53% when compared with Zhang heuristic.

Silva-Filho, et al., with basis on cache tuning parameters, proposes TECH-CYCLES [4] and TEMGA [7] heuristics, the last based on genetic algorithm. An average reduction for instruction caches of 41% and of 15% for data caches in energy usage has been observed to TECH-CYCLES and TEMGA heuristics respectively.

To the best of our knowledge, no previous work has applied Design of Experiments (DoE) [8] for architecture tuning intended for two-level cache hierarchy in order to reduce energy consumption.

3. PROPOSED APPROACH

One of the main concepts introduced in this work is the use of a statistical methodology of data analysis to get a previous knowledge of the application, allowing building an architecture exploration mechanism with a knowledge aggregate value based on result of the obtained analysis. This approach allows using the statistical acquired knowledge to optimize or to develop exploration mechanisms directed toward to characteristics of each application.

In the proposed work we use the statistics technique of design of experiments to analyze the behavior of the memory architecture for each application. Results for this analysis were applied to TECH-DoE heuristic that has a different architecture exploration strategy from existing exploration heuristics; it is defined by DoE analysis. Thus, TECH-DoE is adapted for each application through analysis performed by DoE.

3.1. Design of Experiments (DoE)

The *Design of Experiments (DoE)* is a statistical methodology that considers previous knowledge obtaining of a multivaried process, with the lesser number of attempts. One of the DoE approaches is the *Factorial Planning*, that aims to analyze the effect of the variables of a process and the interaction between its combinations, allowing verifying which variables has greater impact in the process when they are adjusted. There are some factorials planning variations that define the number of

samples necessary to execute the analysis. In this work was used the factorial planning 2^k, where *k* represents the number of factors, that are the variables of the process. In this planning, each factor varies two levels, represented by a maximum and minimal values, or by signals (+) and (-). The levels of a factor represent the values that they can assume. In the 2^k planning, two levels of each factor are chosen, and the experiments will be guided by combination of all factors. Thus, in a planning involving k factors, it can use the minimum and maximum values of each factor as input for the planning, being necessary 2^k experiments to perform the factorial analysis. Each experiment represents a simulation/ observation of the process, where is possible to observe the values of the input/output variables. The same experiment can be replied in order to improve the experiments errors; however in our approach the experiments are not replied. The statistical model of the factorial planning, for the case of two factors with single response is given by equation 1:

$$y_{ij} = \mu+\tau_i+\beta_j+(\tau\beta)_{ij}+\varepsilon_{ij}\,, \qquad (1)$$

where y_{ij} is the observed response by *i-level* of the factor A and *j-level* of the factor B, μ is the average of the overall results of the observed output variable (in our case, energy consumption), τ_i is the main effect of the factor A in the *i-level*, which represents the effect of the factor A in the level *i* related to output variable, β_j is the main effect of the factor B in the j-level, which represents the effect of factor B in level *j* related to output variable, $(\tau\beta)_{ij}$ is the effect of the interaction between τ_i and β_j, and ε_{ij} is the experimental error.

The proposed work uses a factorial planning of 2^6, that represents a planning of 6 factors, each one varying two levels. The six factors represent the configuration parameters of the two-level memory hierarchy, with separate instruction and data caches for both levels: total cache size, line cache and associativity for first and second levels, represented by S1, L1, A1, S2, L2 and A2, respectively. For the levels of each factor the minimum and maximum values of each configuration parameter had been chosen.

When factorial planning is applied to a set of experiments, is possible to observe the effect of each parameter or interaction between parameters in the response of the process. In this work, the effect of the factors adjustment related to energy consumption obtained by a given configuration running a specific application was analyzed. Figure 1 shows the interaction effects between factors with basis on output variable of energy consumption considering the susan_large application from Mibench benchmark suite [9].

Figure 1 shows the order of importance of the interaction between parameters, where the first set of factors have greater influence (EF, DE, F, C, D, DF, CEF, CDE, CF, BCF, BF, BCD, BD, ... , ACD). This order is used to define the exploration strategy of the TECH-DoE heuristic. This strategy is adapted for each application to be analyzed.

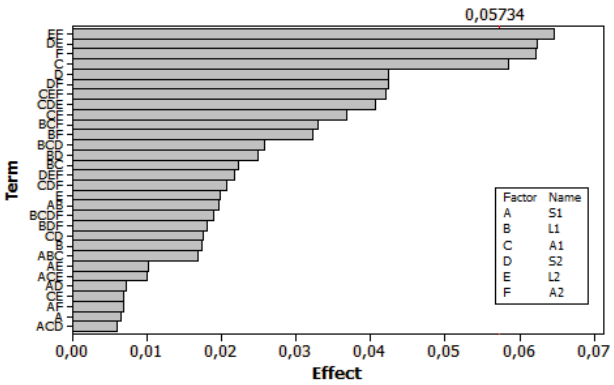

Fig. 1. Pareto Chart of the Effects

3.2. TECH-DoE

TECH-DoE heuristic uses the same setup environment of TECH-CYCLES heuristic to obtain energy consumption (*Simplescalar tool*) [10] and performance statistics (*eCACTI tool*) [11], however with exploration strategy based on planning of experiments results. In this work, we consider similar initial configuration [4] with minimal values for all parameters of the hierarchy (S1, L1, A1, S2, L2 e A2), except to S2 and L2, initialized with the maximum allowed value. The effect order of the interactions between parameters obtained from planning of experiments is used as order of parameters tuning in the TECH-DoE mechanism. For instance, following the analysis showed in Figure 1, "EF" indicates that the simultaneous alteration of the parameters L2 and A2 have greater impact in terms of energy consumption, followed by simultaneous alteration of the parameters S2 and L2, and so one. Thus, the proposed heuristic follows the parameters tuning sequence defined in the factorial planning analysis. The flow chart heuristic functioning can be showed in Figure 2.

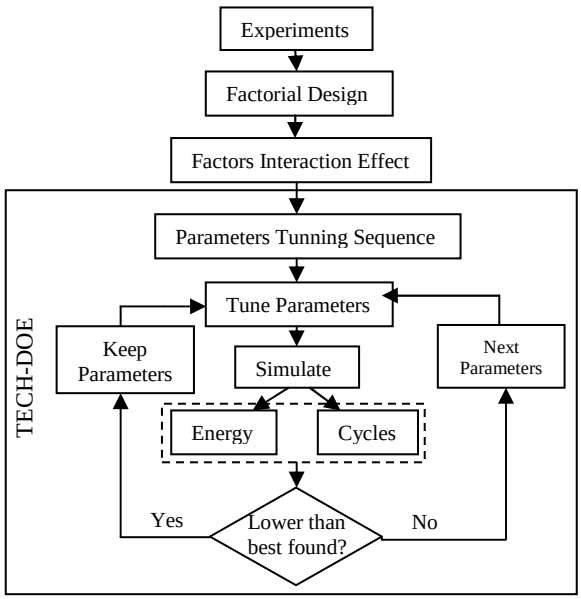

Fig. 2. TECH-DoE exploration mechanism Flow.

Initially, the factorial planning is performed based on a set of experiments, allowing observe the influence degree in the interaction between all factors in terms of energy consumption of the memory system. With basis on the effect order between factors, a parameters tuning sequence is defined and applied by TECH-DoE mechanism, where the parameters with greater impact in terms of energy consumption are firstly tuned in the TECH-DoE mechanism.

While the tuning of one parameter of the sequence obtained from factorial planning will be beneficial for the application, that is, energy consumption and number of cycle's reductions, the tuning in that parameter is kept.

The exploration mechanism continues by varying the same parameter and the procedure stops when the energy consumption or number of cycles of a new cache configuration is bigger than the smallest values obtained by exploration mechanism. After, the next parameter to the sequence is selected and new adjustments in the newly parameter are performed. It is repeated until all parameters of the sequence supplied by experimental planning will be concluded.

4. PRELIMINARY RESULTS

The proposed mechanism was applied for 6 applications from Mibench suite: basicmath_small, bitcount_large, dijkstra_small, dijkstra_large, susan_large and patricia_small. These applications were applied to an architecture composed by a MIPS core processor and a memory hierarchy structure. An input voltage of 1.7V is considered, with write-through scheme and transistor technology of 0.08µm.

Figure 3 illustrates the results from one specific application from Mibench benchmark suite [9] (susan_large), demonstrating the quality of the obtained result. The Minitab tool [12] was used to make the factorial planning 2^6, that requires 64 configurations as samples to factorial analysis, representing about 14% from configurations space for this application. Results for this analysis reveal the parameters with greater impact in terms of energy reduction and the influence order of these parameters was used as parameters tuning sequence for TECH-DoE heuristic.

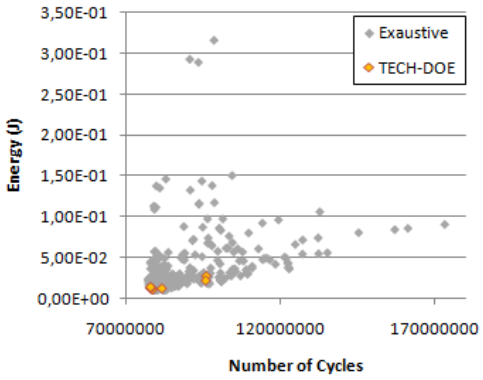

Fig. 3. Exploration using TECH-DoE mechanism.

The complete configuration space (exhaustive approach) for susan_large application is depicted in Figure 3. In this same figure, results of the TECH-DoE mechanism were added in order to demonstrate the difference between it for this application. Results could demonstrate that configurations for TECH-DoE are near to Pareto-Optimal for this application.

In order to validate the proposed approach, the TECH-DoE mechanism has been applied to the selected applications by using the same experimental environment mentioned previously. Our solution was compared to existing exploration mechanism TECH-CYCLES [4] and TEMGA [7], the last based on genetic algorithm. In this analysis, were observed the best configurations reached for each mechanism in terms of energy consumption. A comparative graph for this analysis is showed in Figure 4.

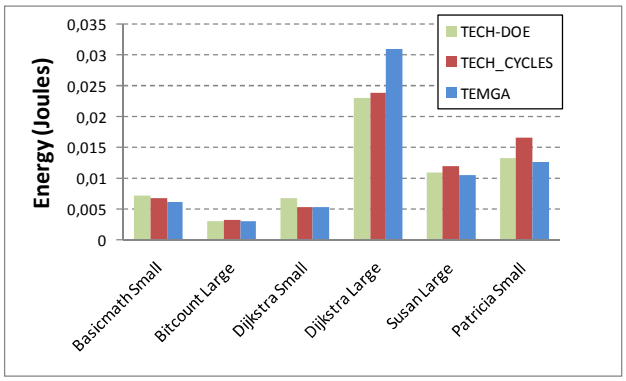

Fig. 4. Best energy consumption reached by configurations found in each mechanism.

Preliminary results in terms of percentage of energy reduction were performed. We got on average, a reduction in terms of energy of approximately 6% for data caches compared to the TECH-CYCLES and TEMGA heuristics. It was observed that TECH-DoE optimization mechanism results are as good as the existing heuristics in terms of energy consumption.

Unlike TECH-CYCLES methodology, DoE approach removes the necessity of understanding of the application behavior in the exploration mechanism. This knowledge is obtained through experiments from DoE approach. Although the using of genetic algorithms also does not requires knowledge, the DoE approach revealed be simpler and with lesser computational cost than similar approach based on genetic algorithm (TEMGA heuristics).

5. CONCLUSIONS

In this work a new exploration mechanism based on design of experiments (DoE) and two-level hierarchies, with separate instruction and data caches for both levels, was proposed. The main idea was to add knowledge to exploration mechanism through statistical analysis of the behavior of experimental samples. This knowledge is used to define the exploration strategy and to optimize the search space for each application.

Analyzing the energy consumption issue for data caches, some preliminary results have been concluded that DoE optimization mechanism reaches an average energy reduction about 6% when compared with existing heuristics. The best configurations found by TECH-DoE were close to the Pareto-Optimal and the energy consumption results were compatible with existing mechanisms.

TECH-DoE also does not require previous knowledge of the application behavior and when compared with similar approach, such as TEMGA heuristics that is based on genetic algorithm, the proposed optimization mechanism has been revealed simpler and with lesser computational cost when compared with the TEMGA heuristic.

We are currently working on to optimize the statistical analysis of the DoE experiments in order to obtain better performance and to extend the analysis for other benchmarks. The DoE-based methodology also can be used in other exploration mechanisms.

6. REFERENCES

[1] Verma, M. and Marwedel, P., *Advanced Memory Optimization Techniques for Low-Power Embedded Processors*, Springer, Netherlands, 2007.

[2] M. Kandemir and A. Choudhary. *Compiler-Directed Scratch Pad Memory Hierarchy Design and Management*, In Proceedings of Design Automation Conference (DAC'02), New Orleans, USA, Jun. 2002.

[3] Zhang, C., Vahid, F., Cache configuration exploration on prototyping platforms. 14th IEEE Interational Workshop on Rapid System Prototyping (June 2003), vol 00, p.164.

[4] A.G. Silva-Filho, F.R. Cordeiro, R.E. Sant'Anna, and M.E. Lima, "Heuristic for Two-Level Cache Hierarchy Exploration Considering Energy Consumption and Performance", In: Integrated Circuit and System Design, Power and Timing Modeling, Optimization and Simulation (PATMOS'06), pp. 75-83, Sep 2006.

[5] A.G. Silva-Filho, C.J.A. Bastos-Filho, D.M.A. Falcão, F.R. Cordeiro, and R.M.C.S. Castro, "An Optimization Mechanism Intended for Two-Level Cache Hierarchy to Improve Energy and Performance using the NSGAII Algorithm". In: International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD'08), pp. 19-26, Oct. 2008.

[6] Gordon-Ross, Ann, Vahid, F., Dutt, Nikil, Automatic Tuning of Two-Level Caches to Embedded Aplications, DATE, pp.208-213 (Feb 2004).

[7] A.G. Silva-Filho, C.J.A. Bastos-Filho, R.M.F. Lima, D.M.A Falcão, F.R. Cordeiro and M.P. Lima. "An Intelligent Mechanism to Explore a Two-Level Cache Hierarchy Considering Energy Consumption and Time Performance", SBAC-PAD, pp 177-184, 2007.

[8] Montgomery, D. C. "Design and Analysis of Experiments", Foruth ed., Wiley, New York, 660p., 1997.

[9] M.R. Guttaus, J.S. Ringenberg, D. Ernst, T.M. Austin, T. Mudge, R.B. Brown, "Mibench: A free, commercially representative embedded benchmark suite". In: IEEE 4th Annual Workshop on Workload Characterization (WWC'01), Ann Arbor, USA, pp.3-14, December 2001.

[10] D. Burger, T.M. Austin, "The SimpleScalar Tool Set, Version 2.0"; Computer Architecture News; Vol 25(3), pp.13-25; June 1997.

[11] N. Dutt, M. Mamidipaka, "eCACTI: An Enhanced Power Estimation Model for On-chip Caches", TR 04-28, Set. 2004.

[12] Factorial Planning Tool, Minitab 1.5. Available in: www.minitab.com, Jun. 06/2009.