A WCET-aware cache coloring technique for reducing interference in real-time systems

Fabien Bouquillon, Clément Ballabriga, Giuseppe Lipari, Smail Niar

Compas'2019 : Parallélisme / Architecture / Système
LIUPPA - IUT de Bayonne, France, du 24 au 28 juin 2019

A WCET-aware cache coloring technique for reducing interference in real-time systems

Fabien Bouquillon[1,2], Clément Ballabriga[1], Giuseppe Lipari[1], Smail Niar[2]

[1] Univ. Lille, CNRS, CentraleLille,UMR 9189 - CRIStAL, Lille, France

[2] Univ. Polytechnique Hauts-de-France, LAMIH/CNRS, Valenciennes, France

Abstract
The time predictability of a system is the condition to give safe and precise bounds on the worst-case execution time of real-time functionalities which are running on it. Commercial off-the-shelf(COTS) processors are increasingly used in embedded systems and contain shared cache memory. This component has a hard predictable behavior because its state depends on the execution history of the systems. To increase the predictability of COTS component we use cache coloring, a technique widely used to partition cache memory. Our main contribution is a WCET aware heuristic which partition task according to the needs of each task. Our experiments are made with CPLEX an ILP solver with random tasks set generated running on preemptive system scheduled with earliest deadline first(EDF).

1. Introduction

Hard real-time systems are found in many different domains, like avionics, automotive, health care services. In such systems, a real-time task has to be executed within predefined timing constraints, whose violation can lead to system failure. Thus, it is important to compute the response time of every task to ensure *a-priori* that it always executes within its time window under all conditions. Schedulability analysis algorithms provide upper bounds on the response time of tasks, which depend upon several parameters such as tasks' execution times and scheduling policy. In turn, execution times depend on the hardware architecture and task's code.
Commercial off-the-shelf (COTS) processors are increasingly used in embedded systems for their low cost and high performance. Most of COTS processors use cache memories to bridge the gap between processor speed and main memory access speed. In particular, cache memories improve performance by reducing the *typical* execution time of a task.
However, in real-time systems we need predictability, that is we need to precisely estimate the *Worst-Case* Execution Time (WCET) of a task. Cache memory state depends on execution history of the system and, its prediction is a challenge due to the increase of tasks running on the system which compete for the same cache memory area. More sophisticated WCET analyses take into account the state of the cache during the execution of the task and provide a tighter WCET. However, these analyses typically assume that every task executes alone in the system, without interference from other tasks. If tasks are executed concurrently and preemptively on

Acknowledgment: This project is supported by «La Fédération de Recherche Transports Terrestres & Mobilité (FR TTM 3733) du CNRS »
Extended version of this work is available at: https://arxiv.org with the name A WCET-aware cache coloring technique for reducing interference in real-time systems

the system, one task may preempt another task and evict cache blocks, making the estimated WCET too optimistic. This type of interference is called *inter-task* interference, as opposed to *intra-task* interference due to a task evicting its own cache blocks.

In the literature, many researchers have been interested in this problem by accounting for the cost of preemption through the so-called *Cache-Related Preemption Delay* [1,8,9]. Another problem arises in multi-core systems with shared cache: a task executing on one processor may evict useful cache blocks for a seconds task executing on a different processor. It is, therefore, necessary to reduce, or eliminate altogether, the inter-task interference caused by the cache conflicts on tasks' execution times.

The goal of this research is to use virtual memory and cache-coloring techniques to reduce inter-task interference: we allocate the virtual pages of a task to physical pages so to minimize conflicts between tasks on set-associative caches. Since cache memory is limited, by doing, so we might increase intra-task conflicts: two pages from the same task may be allocated to two physical pages that correspond to the same position in the cache, thus increasing the task's WCET. We, therefore, propose a methodology to explore the space of possible cache-coloring configurations so to reduce conflicts while maintaining the respect of the timing constraints. We can represent this problem as a variant of the multiple-choice knapsack problem where the colors are the knapsack and the pages the object, but in this variant, the values of the object depend also on the presence of the other objects in the same knapsack. Since the problem's complexity is very large, we propose a combination of Integer-Linear-Programming techniques and heuristics to partition the cache taking into consideration the WCET of each task.

2. Related Works

Predictability of cache memory in real-time systems has been widely explored, especially for the CRPD [1,8,9].

Luniss et al. [8] used simulated annealing to find a code layout in the memory that minimizes the CRPD. However, tasks are not isolated on cache memory, so inter-task and inter-core interference are still present. They used the linker to configure the code layout.

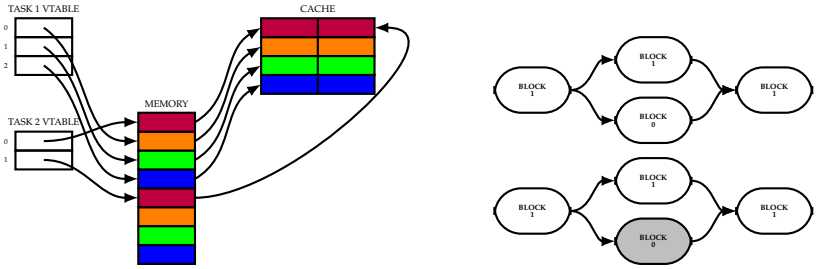
Mancuso et al. [9] propose a complete framework which defines, isolates and locks most important memory areas in memory cache. These techniques are based on cache coloring partitioning and cache locking, its purpose is to reduce conflicts and enhance predictability but the cache is not optimally used because only the most important memories area are in the cache and to access other areas require costly RAM access. In our work, we use their techniques in our heuristics for the pages coloring, but instead of giving all the partitions to the most important data, we reserve a partition to the other data.

Kim et al. [6] propose a practical OS-level cache management scheme using page coloring. They work on partitioned fixed priority preemptive scheduling system where they partition cache memory between cores with page coloring. In their works tasks may share the same cache area, thus intra-core interference is still present.

Ward et al. [10] consider colors as shared resources protected by critical sections, thus priority inversion may occur during execution. To reduce this problem they propose to slice tasks' periods, but their method may force the preempted task to reload its data (the set of data pages that a task may access in one job).

3. System model

In this section, we first present the task model, and then the model of the hardware architecture.



(a) Cache colloring (b) WCET Control Flow Graph

Figure 1: Cache coloring and its impact on WCET

We consider a system of N real-time sporadic tasks $\mathcal{T} = \{\tau_1, \cdots, \tau_N\}$. A task τ_i is an infinite succession of jobs $J_{i,k}(a_{i,k}, c_{i,k}, d_{i,k})$, each one characterized by an arrival time $a_{i,k}$, a computation time $c_{i,k}$ and an absolute deadline $d_{i,k}$. A job $J_{i,k}$ must be executed in the interval of time $[a_{i,k}, d_{i,k}]$, if it misses its deadline then a critical failure occurs.
A sporadic task τ_i can be represented by the tuple (C_i, D_i, T_i, P_i), where C_i represents the worst case execution time (WCET) of the task τ_i ($C_i = \max_{\forall k, k \geq 1}\{c_{i,k}\}$), T_i represents the minimum time between two consecutive arrivals ($T_i \geq \min_k\{a_{i,k+1} - a_{i,k}\}$), D_i is the relative deadline ($\forall k,\ d_{i,k} = a_{i,k} + D_i$), and P_i is the number of distinct virtual pages used by the task.
We consider a set of sporadic tasks with constrained deadlines ($D_i \leq T_i$), scheduled with the preemptive Earliest Deadline First (EDF) scheduler on a single processor. This work can be easily extended to partitioned scheduling on a multi-core system with shared caches. We also assume that tasks are independent, that is they do not share any memory page. We will discuss later how to remove such assumption.
We consider a set-associative cache and denote as $S^{instructions}$ the number of distinct pages that fits into the cache. In this paper we only focus on the *instruction* cache: extension to data cache is the subject of future work. We denote as N^{way} the number of cache way.
The *color* of the j-th virtual page $p_{i,j}$ of task τ_i, denoted as $\kappa_{i,j}$, is an index between $[0, \frac{S^{instructions}}{N^{way}} - (N-1)]$ that denotes the position in the cache where the page will be loaded. Therefore, we search a method for allocating virtual pages to physical pages so that any two different tasks share the minimum possible number of colors (ideally zero).
Main memory size is a multiple of cache memory size which is a multiple of a page size. Therefore, when considering cache coloring at page level, the same page is always mapped to the same cache page(partition of the cache memory of a page size).
Figure 1a represents an example cache coloring technique in a set-associative cache: all pages in main memory with the same color share the same cache page (red color). Thus, the color of each page in main memory can be computed as $\kappa_{i,j} = index(P_{i,j}) mod \frac{S^{instructions}}{N^{way}}$. $\kappa_{i,j}$ depends on the index of the page in main memory, we can use the virtual table of the task to color instructions pages. The configurations of task pages has an impact on the typical execution of the task, thus also on the WCET.

3.1. Recall on WCET analysis

To compute the WCET, we can use measurements or static analysis. Measurements give an optimistic estimation of the WCET because not all inputs and internal states can be tested. The static analysis gives a safer over-estimation of the WCET value. Our static analysis method builds a control flow graph (CFG) of the task and runs various analyses on it (including cache

behavior prediction) to compute an estimation of the WCET. The task's pages allocation has an impact on its WCET: in Figure 1b we show an example of two CFGs of the same task with different page configurations. The node's color represents the color of the page which contains the block, and the edges the possible paths that the execution may take. On the top CFG, Block 1 and Block 0 are not in the same page but use the same area in the cache memory, if the wcet path uses Block 0, then Block 1 will be evicted in cache memory. In the bottom CFG, Block 1 and Block 0 do not use the same cache memory area because the colors of their pages are not the same, thus there will be no eviction and lower execution time.

4. WCET-aware Coloring Heuristics

Our goal is to allocate the virtual memory pages of a set of real-time tasks to the physical memory pages so to minimize the inter-task interference on the cache. In this paper, we try to completely remove the interference by partitioning the cache.
We divide the problem into two steps: 1) At the *macro-level*, we assign a certain number of colors to each task so that the total number of colors is less than or equal to the number of available colors in the cache; 2) At the *micro-level*, for each task separately, with a given number of available colors, we compute the best WCET for that task.
We start by proposing a method for solving the *micro-level*.

4.1. Pages Coloring for a given partition size

Consider that for each possible color combination, it is necessary to perform a WCET analysis. Since that can be very time-consuming, we rule out the complete exploration of all possible combination, and we use a heuristic instead. An overestimation of the number of solutions is given by $(P_i)^{P_i}$ which is exponential.
We consider 2 heuristic algorithms. The first algorithm assigns the same number of pages (approximately) to each color. In particular, if task τ_i is assigned j colors and it has P_i pages, then the same color is assigned to $\lfloor P_i/j \rfloor$ pages. We use a simple modulo: the first $\lfloor P_i/j \rfloor$ pages are assigned to the first color, etc.
The second algorithm classifies pages according to their *importance* in the program. Therefore, we assign each page a *score* that depends on how many times the page is accessed by the program in the Control-Flow-Graph. The score of a page is computed as the sum of the scores of the instructions in the page, and the score of instruction ψ is computed as, $score(\psi) = 10^{l(\psi)}$. Where $l(\psi)$ is the nesting level of loops where the instruction is found: if ψ is not contained in any loop, then $l(\psi) = 0$; if ψ is contained in a loop of first level, then $l(\psi) = 1$; etc. The pages' scores are computed using the OTAWA analysis tool [2]. Then the pages are ordered by decreasing score: if the task τ_i is assigned j colors, the first $j-1$ pages in decreasing score order are assigned a different color, while all other pages are assigned the last remaining color.
Once each page has been assigned a color according to one of the two heuristics above, we launch the OTAWA WCET analysis tool to obtain the corresponding WCET for the task.
We do this for different values of j in the interval $j = [1, S_i^{max}]$, where

$$S_i^{max} = min\left\{P_i, S^{instructions} - (N-1)\right\} \tag{1}$$

and for each value we compute the corresponding WCET $C_i(j)$. These values are used by the ILP solver described in the next section.

4.2. Partitioning cache memory according to the need of the tasks

The distribution of cache memory space can be represented as a Multiple Choice Knapsack Problem(MCKP). In this problem we have a knapsack of limited size and a set of objects of different categories. The problem consists in selecting one and only one object of each category to put in the knapsack.

In our case, the size of the knapsack represents the schedulability constraints; the objective function is the number of colors used (that we want to minimize); the object types are the different tasks, and an object is a configuration of colors for a given task, with the corresponding WCET.

We encode the problem above as an Integer Linear Programming (ILP) problem, and we use CPLEX as a solver. We use the following variables and constraints:

- We define variable $\chi_{i,j} \in \{0,1\}$ to denote the fact that task τ_i has been assigned j colors. Each tasks must be assigned at least one configuration selected: $\sum_{j=1}^{S_i^{max}} \chi_{i,j} = 1$
- The worst case execution time of a task can be expressed as: $C_i = \sum_{j=1}^{S_i^{max}} (C_i(j) \cdot \chi_{i,j})$. where $C_i(j)$ is computed in the micro-level problem.
- We want to minimize the total number of colors used: $\min \sum_{i=1}^{N} \sum_{j=1}^{S^{max}} (j \cdot \chi_{i,j})$ If the value of the objective function for the optimal solution is greater than $\frac{S^{instructions}}{N^{way}}$, then the problem has no feasible solution, and we must resort to other methods for computing the interference (for example by using the CRPD analysis [1]).
- To impose the schedulability of the system, we use the DBF analysis for EDF, first proposed by Baruah [3]. We first represent the utilization constraint $\sum_{i=1}^{N} \frac{C_i}{T_i} \leq 1$ Then we add all inequalities to check that all deadlines are respected. Let $dset(\tau_i) = \{\forall i = 1, \ldots, N, \forall k > 0 | kT_i + D_i \leq DIT\}$. The first definitive IDLE time (DIT) [7], is an instant at which all tasks must complete, and it does not depend on the WCETs of the tasks. Then, we add the following inequalities: $\forall t \in dset(\tau) : \quad \sum_{i=1}^{N} \left(\left\lfloor \frac{t-D_i}{T_i} \right\rfloor + 1 \right) \cdot C_i \leq t$

5. Result

The analysis takes into account a system with a 32 KB set associative memory cache of 2 ways with 512 rows. We consider a page size of 1 KB (this value is defined as a constant in OTAWA and involved timely tool modification if we want to change its value), thus, there are 16 colors available. We test each utilization in the range $[0.30; 1.70]$ (we assume a step of 0.01), with 1000 variation of periods and deadlines of the 8 tasks in Table 3b, taken from well-known standard benchmarks in the literature [4, 5].

First, our method performs a static analysis of each task which gives us a list of $WCET$ according to their number of available colors, the worst of them is selected to compute the periods and deadline with uunifast algorithm ($T_i = WCET_i(worst)/U_i$). To represents constrained deadlines we assign for each task, a deadline in the range of $[WCET_i(worst) + (T_i - WCET_i(worst)) \cdot 0.75) \cdot T_i, T_i]$. In the following figures, the line labeled as *infinite cache* represents the percentage of schedulable tasks set that we can schedule if we have a cache of unbounded size with the WCET list from random coloring.

The *random* line represents the percentage of task schedulable with a random distribution of the cache space between tasks and a random coloring of their pages. Our method (described in the previous section) is represented with the lines labeled *ILP*. The x-axis represents the utilization of the worst distribution with random coloring.

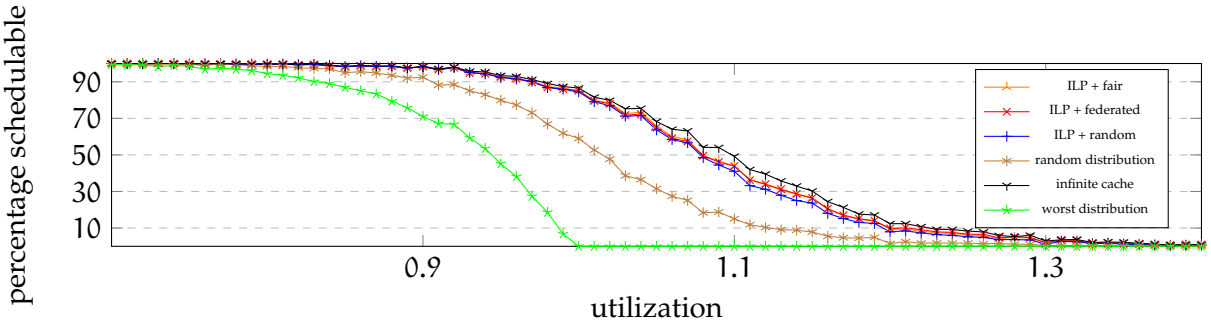
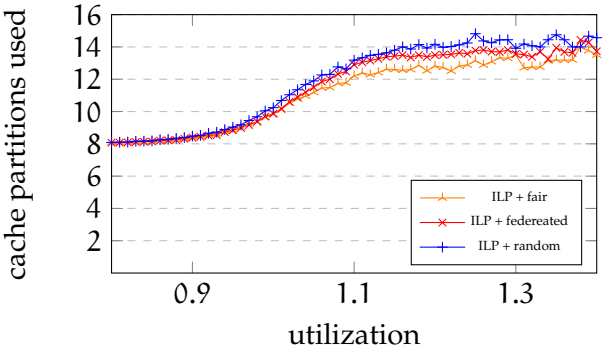


Figure 2: percentage of schedulable task set with constrained deadlines



(a) Average number of cache partitions used with constrained deadlines

Task	pages
compress (Mälardalen)	4
fir (Mälardalen)	2
ndes (Mälardalen)	4
jfdctint (Mälardalen)	3
edn (Mälardalen)	4
crc (Mälardalen)	2
g723_enc (TACLeBench)	8
petrinet (TACLeBench)	8

(b) Tasks used in analysis

Figure 3: performances of heuristics with constrained deadlines

For all heuristics, Figure 2 shows that our method (ILP) increases the amount of schedulable set (more than 20% compared to random distribution for high utilization), but the performances of our heuristics are mitigated compared to the random coloring. On this figure, we do not observe any significant difference between the performance of fair coloring and federated coloring. However, if we look at Figure 3a we can see that fair coloring uses fewer pages than federated and random. So the best heuristics is fair coloring compared to random and federated. This can be explained by the fact that for a low number of colors j, federated coloring isolates only the $j - 1$ most important pages. If the score of the j-th pages is also important, it will experience a significant number of evictions in all the other pages with the lower scores.

6. Conclusion

We proposed an approach based on ILP to partition the cache memory according to the needs of each task for a preemptive system scheduled with EDF. We also propose a heuristic based on our empirical results to find pages layout for each task according to the number of its given colors. Our experimental results confirm an increase of high utilization tasks set schedulable of 20% compared to a random partition of cache memory, however, the performances of our heuristics to coloring task pages are mitigated. We will reduce the granularity to have a method to partition at the granularity of the size of a cache line and explore other heuristics in a future work.

Bibliographie

1. Altmeyer (S.), Davis (R. I.) et Maiza (C.). – Improved cache related pre-emption delay aware response time analysis for fixed priority pre-emptive systems. *Real-Time Systems*, vol. 48, n5, 2012, pp. 499–526.
2. Ballabriga (C.), Cassé (H.), Rochange (C.) et Sainrat (P.). – Otawa: an open toolbox for adaptive wcet analysis. – In *IFIP International Workshop on Software Technolgies for Embedded and Ubiquitous Systems*, pp. 35–46. Springer, 2010.
3. Baruah (S. K.), Mok (A. K.) et Rosier (L. E.). – Preemptively scheduling hard-real-time sporadic tasks on one processor. – In *Real-Time Systems Symposium, 1990. Proceedings., 11th*, pp. 182–190. IEEE, 1990.
4. Falk (H.), Altmeyer (S.), Hellinckx (P.), Lisper (B.), Puffitsch (W.), Rochange (C.), Schoeberl (M.), Sørensen (R. B.), Wägemann (P.) et Wegener (S.). – Taclebench: A benchmark collection to support worst-case execution time research. – In *16th International Workshop on Worst-Case Execution Time Analysis (WCET 2016)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2016.
5. Gustafsson (J.), Betts (A.), Ermedahl (A.) et Lisper (B.). – The Mälardalen WCET benchmarks – past, present and future. – pp. 137–147, Brussels, Belgium, juillet 2010. OCG.
6. Kim (H.), Kandhalu (A.) et Rajkumar (R.). – A coordinated approach for practical os-level cache management in multi-core real-time systems. – In *2013 25th Euromicro Conference on Real-Time Systems*, pp. 80–89. IEEE, 2013.
7. Lipari (G.), George (L.), Bini (E.) et Bertogna (M.). – On the average complexity of the processor demand analysis for earliest deadline scheduling. – In *Proceedings of a conference organized in celebration of Professor Alan Burnss sixtieth birthday*, p. 75, 2013.
8. Lunniss (W.), Altmeyer (S.) et Davis (R. I.). – Optimising task layout to increase schedulability via reduced cache related pre-emption delays. – In *Proceedings of the 20th International Conference on Real-Time and Network Systems*, pp. 161–170. ACM, 2012.
9. Mancuso (R.), Dudko (R.), Betti (E.), Cesati (M.), Caccamo (M.) et Pellizzoni (R.). – Real-time cache management framework for multi-core architectures. – In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2013 IEEE 19th*, pp. 45–54. IEEE, 2013.
10. Ward (B. C.), Herman (J. L.), Kenna (C. J.) et Anderson (J. H.). – Making shared caches more predictable on multicore platforms. – In *2013 25th Euromicro Conference on Real-Time Systems*, pp. 157–167. IEEE, 2013.