Towards a Predictable, High-Performance Instruction Memory Hierarchy in Fixed-Priority Preemptive Real-Time Systems

Abstract

Contemporary computing architectures make an extensive use of cache memories to cope with the increasing speed gap between the processing power of the processing element and the bandwidth of the memory subsystem. In a fixed-priority preemptive real-time system, in addition to obtaining logically correct results, it is also mandatory to satisfy timing constraints. Since the contents of the cache memory depend on the history of the execution of all the tasks on the system, as well as the order in which they were scheduled, this cache-inherent behaviour leads to unpredictable execution times. Therefore, exploiting cache memories in the best way by taking full advantage of its tremendous bandwidth (w.r.t. main memory), is thus a challenging problem for real-time systems designers, since the inclusion of cache memories introduce not so easy to solve problems in the schedulability analysis.

To contribute to the solution of the unpredictability when using instruction cache memories in fixed-priority preemptive real-time systems, this research proposes using an LSM-based memory hierarchy, a novel instruction memory hierarchy that is predictable by design. It offers a guaranteed predictable and repeatable memory latency by providing a hardware mechanism to discriminate which blocks must be loaded into the instruction cache. Furthermore, this automatic, on-demand storage control method minimises the total time that a task must be stalled waiting for the memory by freeing the software from explicit management of the memory hierarchy.

Determinism is achieved in the LSM-based memory hierarchy by protecting and restoring (on-the-fly) instruction cache contents each time there is a context switch. The selection of the instruction cache contents is key to obtain lower worst-case execution/response times. Thus, this dissertation presents sequential and parallel versions of a genetic algorithm, that yield approximate solutions in a competitive amount of computation time.

To examine the validity of both the LSM-based memory hierarchy and the estimated worst-case execution/response times resulting from the instruction cache contents selection by the genetic algorithms, several sets of results have been evaluated. The data in these sets come from several sources and allow assessing the performance of the LSM-based memory hierarchy in front of a locking instruction cache used in a dynamic manner as well as with respect to using a conventional instruction cache. In addition, the results make it possible to contrast the sequential and parallel versions of the proposed genetic algorithm to determine their adequacy in selecting the instruction cache contents.

Keywords: Real-Time Systems, Instruction Cache Memory, Timing Analysis, Worst-Case Execution Time, Worst-Case Response Time, Predictability, Performance, Genetic Algorithms

E. Tamura

Department of Science and Engineering of Computing, Pontificia Universidad Javeriana – Cali