Architectures Session

Heterogeneous Chip Multiprocessor Design

O. Ozturk
Computer Engineering Department,
Bilkent University,
Ankara,Turkey
ozturk@cs.bilkent.edu.tr

Abstract

Chip multiprocessors (CMPs) become an increasingly attractive option for obtaining high performance and low power consumption since it has become a difficult task to obtain more performance out of single-processor designs. As a result, CMPs are widely available in the market and have been used as an attractive option for overcoming the barriers in processor design. As technology scales, the International Technology Roadmap for Semiconductors (ITRS) projects that the number of cores will drastically increase to satisfy performance requirements of future applications.

Despite the many advantages of CMPs over uniprocessor architectures, one of the key questions raised by many researchers is the effectiveness of CMPs. From a hardware point of view, homogeneity inherent to CMPs is a source of limitation in extracting the best utilization from these architectures. Even though complex cores provide higher single-thread performance, they consume more area and power compared to their simpler counterparts. Every application has a different processing need and a memory requirement. Even one single application has different requirements throughout its execution. An application may exploit a high level of instruction-level parallelism where a powerful core will be a better match, whereas a simpler core will suffice for a different application with lower instruction-level parallelism. To overcome the limitations due to the homogeneous behavior of CMPs, Heterogeneous (asymmetric) CMPs have been proposed.

Focusing on such a heterogeneous chip multiprocessor, this paper proposes a heuristic solution to the problem of memory hierarchy management, i.e., the data management across the memory hierarchy. The goal is to determine the optimal data transfers across the different levels of a software-managed multi-level memory hierarchy so that the overall energy consumption in memory is minimized. While our approach operates with a fixed set of processors, in case it detects that a certain processor is not exercised, this can be useful to the designer in refining the processor design as well. Our heuristic approach has two major benefits. First, it can be used for determining optimal data access pattern for a data-intensive application executing on a heterogeneous chip multiprocessor. Second, it can be used as a testbench to evaluate the success of the heterogeneous chip multiprocessor as opposed to homogeneous counterparts.

We are planning to implement the proposed solution within a simulator platform initially and test it using applications. Furthermore, we would like to develop integer linear programming based solutions to compare with the optimum results. We also would like to test our approach on a real heterogeneous chip multiprocessor. Our initial experimental evaluation shows that the proposed approach generates promising results and solution times are not excessive.