



ΠΑΡΟΥΣΙΑΣΗ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

ΗΜΕΡΟΜΗΝΙΑ: Παρασκευή, 15 Απριλίου 2016

ΩΡΑ: 11:00

ΑΙΘΟΥΣΑ: Αίθουσα Σεμιναρίων
Κτήριο Τμήματος Μηχανικών Η/Υ & Πληροφορικής

ΟΜΙΛΗΤΗΣ: Σπυρίδων-Δημήτριος Αγάθος

Θ έ μ α

**«Αποδοτική Υποστήριξη OpenMP για Γενικού Σκοπού
και Ενσωματωμένες Πολυπύρηνες Αρχιτεκτονικές»**

ή

**«Efficient OpenMP Runtime Support for General-Purpose
and Embedded Multi-Core Platforms»**

Επταμελής Εξεταστική Επιτροπή:

- 1. Βασίλειος Β. Δημακόπουλος**, Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων (*επιβλέπων*).
- 2. Παναγιώτα Φατούρου**, Επίκουρη Καθηγήτρια του Τμήματος Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης.
- 3. Γεώργιος Μανής**, Επίκουρος Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.
- 4. Δημήτριος Νικολόπουλος**, Καθηγητής School of Electronics, Electrical Engineering and Computer Science, Queen's University of Belfast.
- 5. Παναγιώτης Χατζηδούκας**, Ερευνητής Γ Computational Science and Engineering Laboratory, ETH Zurich.
- 6. Κωνσταντίνος Μαγκούτης**, Επίκουρος Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.
- 7. Ιωάννης Φούντος**, Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.



Περίληψη

Throughout the history of computers the technological improvements as well as the design and architecture of computing systems were driven by the demand for constantly increasing performance. Nowadays the physical limits of microelectronics have already made their appearance, causing excessive energy consumption and other problems. The answer to these challenges was the introduction of multicore CPUs. Parallel computing became quickly synonymous with mainstream computing. Multicore processors have conquered not only the desktop but also the embedded systems arena while many-core systems are well under way. However, the rise of multicore systems was not accompanied by an analogous improvement in programming models and tools, which would allow easy exploitation of the extra computational resources. The transition from serial to parallel programming models is a laborious procedure and induces many issues.

OpenMP is a very intuitive parallel programming model which can help in dealing with the above issues. OpenMP is the standard programming model for shared memory multiprocessors and is currently expanding its target range beyond such platforms. The tasking model of OpenMP has been used successfully in a wide range of parallel applications. With tasking, OpenMP expanded its applicability beyond loop-level parallelization. Tasking allows efficient expression and management of irregular and dynamic parallelism. Recently, another significant addition to OpenMP was the introduction of device directives that target systems consisting of general-purpose hosts and accelerator devices that may execute portions of a unified application code. OpenMP thus encompasses heterogeneous computing, as well.

This dissertation deals with the problem of designing and implementing a productive and performance-oriented infrastructure to support the OpenMP parallel programming model. The first group of contributions refers to the efficient support of the OpenMP tasking model and its application to provide a novel solution to the problem of nested loop parallelism. We present the design and implementation of a tasking subsystem in the context of the OMPi OpenMP compiler. Portions of this subsystem were re-engineered, and fast work-stealing structures were exploited, resulting a highly efficient implementation of OpenMP tasks for NUMA systems. Then we show how the tasking subsystem can be used to handle difficult problems such as nested loop parallelism. We provide a novel technique, whereby the nested parallel loops can be transparently executed by a single level of threads through the existing tasking subsystem.

The second group of contributions is related to the design and implementation of efficient OpenMP infrastructures for embedded and heterogeneous multicore systems. Here we present the way we enabled OpenMP exploitation of the STHORM accelerator. An innovative feature of our design is the deployment of the OpenMP model both at the host and the fabric sides, in a seamless way. Next we present the first implementation of the OpenMP 4.0 accelerator directives for the Parallella board, a very popular credit-card sized multicore system consisting of a dual-core ARM host processor and a distinct 16-core Epiphany co-processor.



Finally, we propose a novel compilation technique which we term CARS; it features a Compiler-assisted Adaptive Runtime System which results in application-specific support by implementing each time only the required OpenMP functionality. The technique is of general applicability and can lead to dramatic reduction in executable sizes and/or execution times.