QUAD: A Sophisticated Memory Access Profiling Toolset

S. Arash Ostadzadeh, *Student Member, IEEE,* and Koen Bertels, *Member, IEEE*

Abstract—Computer manufacturers have already embarked on the heterogeneous multi-processors roadmap due to the saturation of single-processor performance scaling. Integrating several cores on a chip or using reconfigurable devices as processing elements, is already into practice, and many-core architectures are on the horizon. The shift to an increasing number of cores and heterogeneous architectures has placed new burdens on the programming community. Traditionally, applications have been developed considering a single processor, and they need to be parallelized and/or optimized to take advantage of the new breed of heterogeneous multi-/many-core systems. This requires programmers to understand how to design, write and debug parallel applications effectively.

The increased complexity of programming on multi-processors platforms requires more insight into program behavior. It also necessitates the use of tools that can support programmers in migrating existing applications to these emerging platforms, as well as, in writing new applications. Programmers need increasingly sophisticated methods for profiling, instrumentation, measurement, analysis, and modeling of applications. Particularly, tools to thoroughly investigate the memory access behavior of applications become crucial due to the widening gap between the memory bandwidth/latency compared to the processing performance. Toward this, we developed a dynamic profiling framework that focuses on the extraction of memory access related information during the execution of an application. We present the current status of QUAD, a sophisticated toolset that provides a detailed overview of the runtime behavior of the memory access pattern of an application. The toolset consists of several modules, each examining the memory access behavior of an application with a certain objective. The basic QUAD module is used to detect the actual data dependencies between a pair of communicating functions. There are several extensions for the main module which implement added features. tQUAD reveals the memory bandwidth usage of each function in an application in terms of relative execution timings. xQUAD is able to provide a global view of memory access behavior of an application and a detailed per function information on the source code data object granularity. cQUAD tracks all the data communication events between a pair of cooperating functions to expose the overall communication pattern between the functions. The extracted information by the QUAD toolset is critical for porting an application onto a heterogeneous reconfigurable platform, as well as, for code optimizations and parallel application development support.

Index Terms—Heterogeneous Multi-processors, Reconfigurable Computing, Dynamic Profiling, Instrumentation, Program Analysis, Performance Evaluation.

- *S.A. Ostadzadeh and K. Bertels are with the Department of Computer Engineering, Delft University of Technology, Delft, the Netherlands.*
  E-mail: {S.A.Ostadzadeh, K.L.M.Bertels}@TUDelft.nl