A configurable and scalable multi-core architecture template supporting hybrid model of computation

Giuseppe Tuveri, Sebastiano Pomata, Simone Secchi, Paolo Meloni

{giuseppe.tuveri, sebastiano.pomata, simone.secchi, paolo.meloni} @diee.unica.it

DIEE - Dept. of Electrical and Electronic Engineering, University of Cagliari, P.zza d’Armi, 09123 Cagliari, Italy

Abstract

The poster presents a customizable NoC-based multi-core architectural template supporting hybrid model-of-computation, exploiting simultaneously message passing and shared address space. The template is conceived as a composition of configurable modules; architectural parameters such as number and interface of cores, network topology and features and memory structure organization can be changed at design time. The network interfaces have been enriched to support both the communication models, enabling memory-to-memory transfers, initiated by the processors programming memory-mapped circuitry. Hardware event files, readable and modifiable via software, have been developed and are exploited for synchronization purposes. The developed hardware architectures are depicted in the poster together with the results of their implementation on a state-of-the-art technology library. APIs and related implementations for the main communication primitives of both models of communication have been developed and are provided with the platform. Further results presented in the poster consist in an assessment of the performance obtainable using the hybrid approach. Evaluations are provided for single synchronization actions (identified actions are task dispatching, point-to-point or all-to-all synchronization, data dispatching) and at application level (several video/image processing and scientific calculation kernels are used as test applications). Template configurations are automatically implementable on FPGA to enable rapid prototyping.
A configurable and scalable multi-core architecture template supporting hybrid model of computation

G. Tuveri, S. Pomata, S. Secchi and P. Meloni
giuseppe.tuveri@di.ee.unica.it

Abstract
The poster presents a customizable NoC-based multi-core architectural template supporting hybrid model of computation, exploiting simultaneously message passing and shared address space. The template is conceived as a composition of configurable modules; architectural parameters such as number and interface of cores, network topology and features and memory structure organization can be changed at design time. The network interfaces have been enriched to support both the communication models, enabling memory-to-memory transfers, initiated by the processors communicating memory-mapped circuitry. Hardware event files, readable and modifiable via software, have been developed and are exploited for synchronization purposes. The developed hardware architectures are depicted in the poster together with the results of their implementation on a state-of-the-art technology library. APIs and related implementations for the main communication primitives of both models of computation have been developed and are provided with the platform. Template configurations are automatically implementable on FPGA to enable rapid prototyping.

Motivation
State-of-the-art multi-core architectures often require heterogeneity to be efficiently tailored for a given application or application domain. This involves not only the processing elements or the communication infrastructure. The model used as reference by the programmer has also a deep impact on the performances. Once the system topology is chosen, different synchronization actions may perform differently when implemented referring to SM or MP model of computation. Here we report some examples.

Architectural template overview

Library of Hardware Abstraction APIs

The usability of the proposed architectural template and components is improved by means of a set of APIs that hide the hardware implementation details to the programmer. APIs implement the main synchronization actions in both supported models of computation, allowing the use of a hybrid programming model.

Performance trade-off

Experimental results show that, once the architecture configuration is selected (plotted results refer to a homogeneous mesh-like topology including 4 processors), different synchronization and communication actions perform differently when implemented referring to SM or MP model of computation. Here we report some examples.

Implementation results

The overhead in terms of hardware costs related to the modules providing support for synchronization is dependent on their configuration chosen at design time. Usual configurations have hardware costs that can be considered affordable when compared with the other system components (results refer to a 90 nm technology library).

FPGA-based prototyping framework

- A “topology builder” (translating very high-level directives into a fast and accurate prototyping platform)
- Support for performance extraction during the prototyping (modalities and granularity specified already at system level)
- Support for technology awareness (back-annotation of the prototyping data with power/frequency models)