

Deeply Heterogeneous Many-Accelerator Infrastructure for HPC Architecture Exploration

J. Flich,^{*} A. Cilardo,^{†§} M. Kovač,[‡] R. Tornero,^{*} M. Gagliardi,[†]
E. Fusella,[‡] J. M. Martínez,^{*} T. Picornell^{*}

^{*}Universitat Politècnica de València, Spain

[†]University of Naples Federico II, Italy

[§]Centro Regionale ICT (CeRICT), Italy

[‡]University of Zagreb, Croatia

^{*}jfflich@disca.upv.es, [†]acilardo@unina.it, [‡]mario.kovac@fer.hr

Keywords Many-core accelerators, FPGA-based computing, Interconnection networks

Topic 1 – Architectures

1 Abstract

It is well known that heterogeneous computing is being adopted in HPC systems, with the adoption of GPGPUs and manycore accelerators, and currently with the early adoption of FPGA systems. The FETHPC MANGO project emerged to analyze future heterogeneous manycore architectures for HPC. In this paper we describe the details of the heterogeneous architecture proposed in the MANGO project. The paper focuses mainly on the description of the interconnect design, the accelerators layout strategy, and the design of manycore accelerators for specific target applications within MANGO. We conclude the paper with the description of one of the targeted applications within MANGO.

2 Introduction

Heterogeneous computing is becoming a trend in HPC systems. In fact, with the adoption of GPGPUs more power-efficient computation can be achieved for specific applications. As a consequence, manycore accelerators running as slave processors have attracted attention and have ended

up being included in large systems. More recently, FPGA-based systems also got significant attention for very specific application scenarios. This trend is backed up by the running systems listed in the Green 500 ¹ and Top 500 ² lists. In addition, new emerging applications become candidates to be run on HPC systems, as HPC is now seen as a valuable tool beyond the traditional application domains of oil & gas, finance, meteorology, and scientific computation.

In this context, the MANGO project [1] emerges as a tool for the extensive exploration of deeply heterogeneous systems composed of various types of accelerators and high-end computers. The project is building an HPC infrastructure for the rapid prototyping and evaluation of many-core architectures.

The system is compounded of two main components sharing a common rack infrastructure: on one hand, high-end servers with GPGPU support building the so-called General purpose Nodes (GN) and, on the other hand, a set of motherboards with built-in high-end FPGAs and Zynq boards all interconnected pin-to-pin. This is viewed as the Heterogeneous Nodes (HN). In this paper we describe the details of the HN architecture, mainly focusing on the interconnect design, the accelerators layout strategy, and the design of manycore accelerators for specific target applications within MANGO. We conclude the paper with the description of one of the targeted applications within MANGO.

¹<https://www.top500.org/green500/lists/2016/11/>

²<https://www.top500.org/lists/2016/11/>