Exploring Manycore Architectures for Next-Generation HPC Systems through the MANGO Approach

José Flich, Giovanni Agosta, Philipp Ampletzer,
David Atienza Alonso, Carlo Brandolese, Etienne Cappe,
Alessandro Cilardo, Leon Dragić, Alexandre Dray, Alen Duspara,
William Fornaciari, Edoardo Fusella, Mirko Gagliardi,
Gerald Guillaume, Daniel Hofman, Ynse Hoornenborg,
Arman Iranfar, Mario Kovač, Simone Libutti, Bruno Maitre,
José Maria Martínez, Giuseppe Massari, Koen Meinds,
Hrvoje Mlinarić, Ermis Papastefanakis, Tomás Picornell, Igor Piljić,
Anna Pupykina, Federico Reghenzani, Isabelle Staub,
Rafael Tornero, Michele Zanella, Marina Zapater, Davide Zoni

PII: S0141-9331(18)30024-3
DOI: 10.1016/j.micpro.2018.05.011
Reference: MICPRO 2692

To appear in: *Microprocessors and Microsystems*

Received date: 17 January 2018
Revised date: 2 May 2018
Accepted date: 15 May 2018

Please cite this article as: José Flich, Giovanni Agosta, Philipp Ampletzer, David Atienza Alonso, Carlo Brandolese, Etienne Cappe, Alessandro Cilardo, Leon Dragić, Alexandre Dray, Alen Duspara, William Fornaciari, Edoardo Fusella, Mirko Gagliardi, Gerald Guillaume, Daniel Hofman, Ynse Hoornenborg, Arman Iranfar, Mario Kovač, Simone Libutti, Bruno Maitre, José Maria Martínez, Giuseppe Massari, Koen Meinds, Hrvoje Mlinarić, Ermis Papastefanakis, Tomás Picornell, Igor Piljić, Anna Pupykina, Federico Reghenzani, Isabelle Staub, Rafael Tornero, Michele Zanella, Marina Zapater, Davide Zoni, Exploring Manycore Architectures for Next-Generation HPC Systems through the MANGO Approach, *Microprocessors and Microsystems* (2018), doi: 10.1016/j.micpro.2018.05.011

Exploring Manycore Architectures for Next-Generation HPC Systems through the MANGO Approach

José Flich[a,*], Giovanni Agosta[b,*], Philipp Ampletzer[c], David Atienza Alonso[d], Carlo Brandolese[b], Etienne Cappe[i], Alessandro Cilardo[e], Leon Dragić[f], Alexandre Dray[g], Alen Duspara[f], William Fornaciari[b], Edoardo Fusella[j], Mirko Gagliardi[j], Gerald Guillaume[g], Daniel Hofman[f], Ynse Hoornenborg[h], Arman Iranfar[d], Mario Kovač[f], Simone Libutti[b], Bruno Maitre[i], José Maria Martínez[a], Giuseppe Massari[b], Koen Meinds[h], Hrvoje Mlinarić[f], Ermis Papastefanakis[i], Tomás Picornell[a], Igor Piljić[f], Anna Pupykina[b], Federico Reghenzani[b], Isabelle Staub[g], Rafael Tornero[a], Michele Zanella[b], Marina Zapater[a], Davide Zoni[b]

[a] *Universitat Politècnica de València, Spain*
[b] *DEIB – Politecnico di Milano, Italy*
[c] *PRO DESIGN Electronic GmbH, Germany (provider of FPGA-based Hardware - www.profpga-hpc.com)*
[d] *ESL – École Politechnique Fédérale de Lausanne (EPFL), Switzerland*
[e] *Centro Regionale Information Communication Technology SCRL, Italy*
[f] *University of Zagreb, Croatia*
[g] *Eaton Industries SAS, France*
[h] *Philips Medical Systems, The Netherlands*
[i] *Thales Communications & Security, France*
[j] *Universita' degli Studi di Napoli Federico II, Italy*

1. Introduction

The push towards Exascale is going to radically change High-Performance Computing (HPC). First, the sheer amount of computational resources available are pushing the energy envelope available through the power grid to the point where the size of an HPC centre may be constrained by the availability of power supply. Second, the increase in scale of HPC resources across the world is enabling new use case scenarios, where players previously unable to access HPC resources may now do so through innovation in delivery modes, e.g. through cloud HPC [1]. Thus, the evolution of HPC hardware and software architectures needs to embrace technologies with high performances and low power consumptions. The current trend is to leverage application-based customization to this end. Deeply heterogeneous architectures can provide such performance/watt improvements, but are clearly much more difficult to program and manage. Furthermore, new application classes, that are QoS sensitive, are entering the HPC domain. In particular, applications such as video transcoding or medical imaging need time-predictability. Since time-predictability and QoS are often not taken into account in HPC, it is mandatory to extend the traditional optimization space from power/performance to *power*, *performance*, and *predictability* – the PPP space. In fact, predictability, power, and performance appear to be three inherently diverging perspectives on HPC.

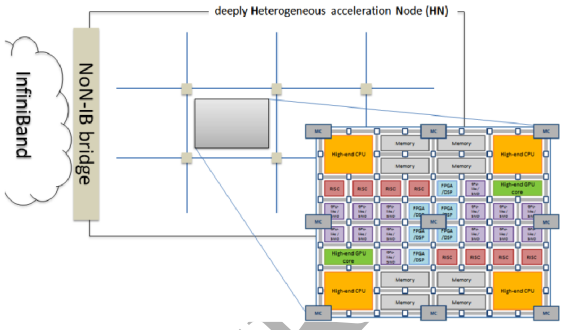

Figure 1: MANGO Hardware Architecture

MANGO's [2, 3] key goal consists in addressing the PPP space by achieving extreme resource efficiency in future QoS-sensitive HPC. The present research investigates the architectural implications of HPC applications' requirements to define a new generation of high-performance, power-efficient, deeply heterogeneous architectures with native mechanisms for isolation and QoS.

1.1. The MANGO Approach

Currently, the major challenge faced by HPC is the performance/power efficiency. Looking straight at the heart of the problem, the hurdle to the full exploitation of today's computing technologies ultimately lies in the gap between the applications' demand and the underlying computing architecture: the better the match between

*Corresponding author