Computer Architecture News

A Publication of the
Association for Computing Machinery
Special Interest Group on Computer Architecture

Vol. 31, No. 1 – March 2003

MEDEA Workshop Guests Editor's Introduction

S. Bartolini*, P. Foglia+, C.A. Prete+

*Department of Information Engineering, University of Siena, Siena, Italy
+ Department of Information Engineering, University of Pisa, Pisa, Italy
(s.bartolini, foglia, prete)@iet.unipi.it

In this Issue of ACM SigArch Newsletter, we present eight papers form the MEDEA Workshop, jointly held with the International PACT (Parallel Architecture and Compilation Techniques) 2002 Conference [1], [2].

The MEDEA Workshop wants to be a forum where people from academy and industry meet, discuss and exchange their research ideas and experiences on relevant and promising computer architecture themes. Following the high level of interest in the first two MEDEA workshops, in the 2002 edition, MEDEA focusses on On-Chip Multiprocessor, multithreading, latency reducing and tolerating techniques and their applications to general purpose and embedded systems. Previous editions of MEDEA focussed on the Memory Decoupling Architecture Concepts from the original idea of J. E. Smith [3] in the context of advanced processor architectures such as Superscalar and VLIW.

Exploiting Instruction Level Parallelism, Out-of-Order Execution, Branch Prediction are common techniques utilized to enhance the performance of microprocessors, but memory latencies still remain the major sources of performance degradation. As observed also in previous edition of MEDEA [4], modern processors can tolerate latencies of only 10 cycles, while the processor-memory gap is going to exceed 100 cycles, and the gap is even more dramatic in multiprocessor systems. On Chip-Multiprocessors reduce memory latencies integrating and optimizing memory hierarchies on a chip, and try to take benefit from Thread Level Parallelism. Multithreaded Architectures hide long latencies operations (from memory or FU) by switching and/or fetching instructions from different thread of execution. As addressed by PACT-2002 keynote speakers, multithreading and multiprocessing on chip are becoming key design issues of today microprocessor generations. They will bring higher levels of parallelism in even mainstream server platforms, and compiler writers and application developers may utilize this parallelism to further speed-up the performance. So, it is a propitious time to investigate the potential of new architectures and solutions in the field.

We will briefly introduce the contributions presented in this issue.

In the paper "Fresh Breeze: A Multiprocessor Chip Architecture Guided by Modular Programming Principles", Jack B. Dennis presents his project concerning the architecture and design of a multiprocessor chip that can achieve high performance while honouring six important principles for supporting modular software construction. The proposed Fresh Breeze chip will incorporate three basic ideas: simultaneous multithreading, a shared 64-bit address space, which can abolish the conventional distinction between ''memory'' and the file system, and write-only

memory chunks with hardware garbage collector that can eliminate the need of coherence protocols.

Morano et. al describe a microarchitecture that achieves high performance on conventional single-threaded programs without compiler assistance. The proposed approach pushes towards a drastic increase in the number of the simultaneously in-flight instructions to improve ILP. The authors present a basic overview of their microarchitecture as a large mesh of processing elements and discuss how they address the problems associated with such microarchitecture: scalability, control flow and memory latency. The microarchitecture makes use of control and value speculative, multi-path, and out-of-order execution. The authors explore several geometries and discuss various design tradeoffs.

Almasi et Al. present Cyclops, a new family of multithreaded architectures, which integrates processing logic, main memory and communications hardware on a single chip. They evaluates several alternative Cyclops designs with different relative costs and trade-offs finding that, by increasing the number of threads sharing a floating point unit, the Cyclops architecture can hide fairly high cache and memory latencies.

Mohamed M. Zahran investigates the performance of some known cache memory hierarchy design applied to single chip multiprocessors. He provides indications about suitable and to be avoided configurations for CMP with coherent cache.

The paper by Gary Grewal and Thomas Wilson presents a methodology, based on an Enhanced Genetic Algorithm (EGA), for assigning data objects to dual-bank memories. The EGA has been incorporated into a retargetable, optimizing compiler for embedded systems, currently under development at the University of Guelph.

Ramacher et al. report the development, at Infineon Technologies AG, of a fully programmable vision processor, addressing the requirements of low power, low costs, and high computational performance typical of such environment. The prototype achieves a peak 100 GOPS with a 0.13 μm CMOS technology, and a power consumption of less than 500 mW.

The paper by Pitsianis and Pechanek discuss the concept of indirect very long instruction word (iVLIW) architecture and its implementation on the BOPS ManArray family of multiprocessor digital signal processors (DSP). They present and approach with small caches of VLIWs localized in each processing element. Experimental results demonstrate the effectiveness of the approach.

Shimizu and Takatori present the implementation of a Super Page Kernel for the Linux operating system on Alpha, Sparc64 and IA32 platforms. The bigger pages allow the applications to reach a wider working set without generating TLB misses. The authors show that with super page kernel, a matrix transpose program can run up to 4 times faster, and the SPEC CPU numbers can be 10% higher than normal kernel on Alpha. The authors show improvements in performance of the Super Page Kernel on a ES-45 Alpha multiprocessor and highlight the applicability of the technique to a simplified CMP system.

Finally, we host an invited paper by Bechini et al., in which they report the experience in designing a single chip multiprocessor for a cartographic dedicated system. In

particular, they adopt symmetrical bi-processor architecture, and analyze the choice of the coherence strategy and the optimization of the system.

We wish to thank the MEDEA Program Committee (Fumio Arakawa, SuperH Japan Ltd., Tokyo, Japan; Alessio Bechini, University of Pisa, Italy; Ricardo Bianchini, Rutgers University, NJ, USA; Doug Burger, University of Texas at Austin, TX, USA; Mats Brorsson, Royal Institute of Technology in Stockholm, Sweden; Roberto Giorgi, University of Siena, Italy; Antonio Gonzalez, Universitat Politècnica de Catalunya, Spain; Ali Hurson, Pennsylvania State University, PA, USA; Liviu Iftode, Maryland University, NJ, USA; David Kaeli, Northeastern University, MA, USA; Krishna Kavi, University of North Texas, TX, USA; Stephen Keckler, University of Texas at Austin, TX, USA; Avi Mendelson, Intel, Israel; Aleksander Milenkovic, University of Alabama In Huntsville, AL, USA; Veljko Milutinovic, University of Belgrade, Serbia, YU; Sanjay Patel, University of Illinois, Il, USA; Jelica Protic, University of Belgrade, Serbia, YU; Alan J. Smith, Univ. of California, Berkeley, CA, USA; Jared Stark, Intel, USA; Theo Ungerer, University of Augsburg, DE, EU; Mateo Valero, Universitat Politècnica de Catalunya, Spain) for their excellent referring work that helped increased the quality of the accepted papers.
We want to thank also G. Lettieri, F. Marcelloni and L. Rizzo, from the University of Pisa, for their precious help in the reviewing process.
Finally, a special thank to Roberto Giorgi and Jelica Protic, organizers of the previous edition of MEDEA, for their invaluable help and support.

References

[1] PACT-2002 Conference, http://moss.csc.ncsu.edu/pact02/
[2] PACT Conferences, http://www.pactconf.org
[3] J. E. Smith, "Decoupled Access/Execute Architectures", Proceedings of 9th International Symposium on Computer Architecture, pp. 112-119, May 1982.
[4] R. Giorgi, "Memory Decoupled Architectures and related Issues Guest Editor's Introduction", IEEE TCCA Newsletter, January 2001