

Seventh Framework Programme – Theme 3.6: Computing Systems

Grant agreement n°248776



D6.3

“Publication Report – Intermediate”

WP6: “Exploitation & Dissemination”

Author: Christian FABRE, CEA.

Date: Tuesday 13th March, 2012

1	CEA (Coord.)	Commissariat à l'énergie atomique et aux énergies alternatives, Laboratoire d'électronique et des technologies de l'information
2	UJF	Université Joseph Fourier Grenoble 1 – VERIMAG
3	ETHZ	Eidgenössische Technische Hochschule Zürich
4	UNIBO	Università di Bologna
5	STM	STMicroelectronics
6	EPFL	École polytechnique Fédérale de Lausanne

Version: v1.1 – Status: Public

CEA ref.: LETI/DACLE/12-0202

Versions of the Document

Version	Date	Author	Comment
0.1	June 13th, 2011	C. FABRE	Initial Version
0.8	July 8th, 2011	C. FABRE	All publis included
0.9	July 11th, 2011	C. FABRE	Added statistics. Final review by partners.
1.0	July 12th, 2011	C. FABRE	Delivered.
1.1	March 12th, 2012	C. FABRE	Public version.

Contributors

Christian FABRE (CEA)

© 2011 CEA for PRO3D

Contents

1	Introduction	4
2	Publication Accepted from January 2010 until June 2011	4
2.1	Publications per Date	4
2.2	Publications per Kind	4
2.3	Conferences & Workshop per Location	4
	References	5
	Scientific Contributions Published During the 1st Reporting Period (2010)	5
	Scientific Contributions Published During the 2nd Reporting Period (2011)	11

1 Introduction

The objectives of this report is to provide a summary of PRO3D publications for the first 18 months of the project.

Section 2 below provides some statistics, while each publication, with its abstract, is presented individually in the remaining of this report.

2 Publication Accepted from January 2010 until June 2011

As of July 11th, 2011, PRO3D had 40 accepted publications.

2.1 Publications per Date

The publication breakdown per year is the following:

- 19 publications in 2010: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19];
- 21 publication for 2011: [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40].

2.2 Publications per Kind

The publication breakdown per kind is the following:

- 30 contributions to international conferences: [1, 2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 15, 16, 17, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 35, 36, 39, 40];
- 5 contributions to international workshops: [8, 18, 19, 33, 34];
- 5 articles in international journals [10, 14, 27, 37, 38];

2.3 Conferences & Workshop per Location

The breakdown of the 35 PRO3D contributions to events (conferences or workshop) is the following:

- 15 contributions to events hosted in Europe;
- 12 contributions to events hosted in North America;
- 3 contributions to events hosted in South Americas;
- 5 contributions to events hosted in Asia.

Scientific Contributions Published During the 1st Reporting Period (2010)

- [1] “Automated Conflict-free Distributed Implementation of Component-Based Models”. B. Bonakdarpour, M. Bozga, M. Jaber, J. Quilbeuf and J. Sifakis In Proceedings of IEEE Fifth International Symposium on Industrial Embedded Systems - SIES 2010, University of Trento, Italy, July 7-9, 2010, pp. 108–117.

We propose a method for generating distributed implementations from high-level models expressed in terms of a set of components glued by rendezvous interactions. The method is a 2-phase transformation preserving all functional properties. The first phase is a source-to-source transformation from global state to a partial state model (to relax atomicity). This transformation replaces multi-party rendezvous interactions by `send()/receive()` primitives managed by a set of automatically generated distributed schedulers. These schedulers are conflict-free by construction in the sense that they do not require communication in order to safely execute interactions of the high-level model. In the second phase, from the transformed model in phase one, we generate C++ distributed code using either TCP sockets or MPI to implement `send()/receive()` primitives. Our method is fully implemented in a tool for automatic generation of distributed applications. We present experimental results using different case studies.

- [2] “From High-Level Component-Based Models to Distributed Implementations”. B. Bonakdarpour, M. Bozga, M. Jaber, J. Quilbeuf & J. Sifakis. In Proceedings of the 10th International conference on Embedded software, EMSOFT 2010, Scottsdale, Arizona, USA, October 24-29, 2010, pp. 209–218.

Although distributed systems are widely used nowadays, their implementation and deployment is still a time-consuming, error-prone, and hardly predictive task. In this paper, we propose a methodology for producing automatically efficient and correct-by-construction distributed implementations by starting from a high-level model of the application software in BIP. BIP (Behavior, Interaction, Priority) is a component-based framework with formal semantics that rely on multi-party interactions for synchronizing components and dynamic priorities for scheduling between Interactions. Our methodology transforms arbitrary BIP models into Send/Receive BIP models, directly implementable on distributed execution platforms. The transformation consists of (1) breaking atomicity of actions in atomic components by replacing strong synchronizations with asynchronous `send()/receive()` interactions; (2) inserting several distributed controllers that coordinate execution of interactions according to a user-defined partition, and (3) augmenting the model with a distributed algorithm for handling conflicts between controllers. The obtained Send/Receive BIP models are proven observationally equivalent to the initial models. Hence, all the functional properties are preserved by construction in the implementation. Moreover, Send/Receive BIP models can be used to automatically derive distributed implementations. Currently, it is possible to generate stand-alone C++ implementations using either TCP sockets for conventional communication, or MPI implementation, for deployment on multicore platforms. This method is fully implemented. We report concrete results obtained under different scenarios (i.e., partitioning of the interactions and choice of algorithm for distributed conflict resolution).

- [3] “Methods for Knowledge Based Controlling of Distributed Systems”. S. Bensalem, M. Bozga, S. Graf, D. Peled & S. Quinton. International Symposium on Automated Technology for Verification and Analysis, ATVA 2010, September 2010, Singapore, LNCS 6252, pages 52-66.

Controlling concurrent systems to impose some global invariant, is an undecidable problem. One can gain decidability at the expense of reducing concurrency. Even under this flexible design assumption,

the synthesis problem remains highly intractable. One practical method for designing controllers is based on checking knowledge properties upon which the processes can make their decisions whether to allow or block transitions. A major deficiency of this synthesis method lies in calculating the knowledge based on the system that we want to control, and not on the resulted system. The original system has less knowledge, and as a result, we may introduce far more synchronization than needed. In this paper we show techniques to reduce this overhead.

- [4] “*Integrating Architectural Constraints in Application Software Using Model Transformations in BIP*” Ananda Basu, Saddek Bensalem, Paraskevas Bourgos, Marius Bozga & Joseph Sifakis. Invited presentation at IEEE International High Level Design Validation and Test Workshop. Anaheim, USA. June 10th-12th 2010.

Performance of embedded applications strongly depends on features of the hardware platform on which they are deployed. A grand challenge in the design of complex embedded systems is developing methods and tools for modeling and analyzing the behavior of an application software running on a given hardware architecture. We propose a rigorous method that allows to obtain a model which faithfully represents the behavior of a mixed hardware/software system from a model of its application software and a model of its underlying hardware architecture. The method takes a model of the application software in BIP, a model of the hardware architecture in XML and a mapping associating read and write operations of the application software with execution paths in the architecture. It builds a model of the corresponding system in BIP. The latter can be simulated and analyzed for the verification of both functional and extra-functional properties. The method consists in progressively enriching the application software model. It involves two steps: (1) The identification of the a set of hardware components used in the system model. These are included in the initial model. They are obtained by a systematic decomposition of the architecture through an abstract grammar. (2). The application of a sequence of source-to-source transformations to synthesize run-time routines implementing software channels. The transformations are correct-by-construction. In particular they preserve functional properties of the application software. We have identified a minimal and complete set of transformation rules, that are proved to be correct. The system model is highly parametrized and allows exible integration of specic target architecture features, such as bus policy and scheduling policy of the processor cores The method has been implemented for software applications and hardware architectures described in the DOL tool for performance evaluation. We plan to test benchmark applications on a simplied MARM architecture.

- [5] “*Thermal-Aware Compilation for System-on-Chip Processing Architectures*”. Aly, Mohamed M. Sabry ; Rodrigo, Ayala ; Luis, José ; Atienza Alonso, David. Proceedings of the 20th ACM Great Lakes Symposium on VLSI (GLSVLSI 2010), vol. 1, num. 1, 2010, p. 221–226, New York: ACM Press. ISBN: 978-1-4503-0012-4/10/06. Providence, Rhode Island, USA. May 16-18 2010.

The development of compiler-based mechanisms to reduce the percentage of hotspots and optimize the thermal profile of large register files has become an important issue. Thermal hotspots have been known to cause severe reliability issues, while the thermal profile of the devices is also related to the leakage power consumption and the cooling cost. In this paper we propose several compilation techniques that, based on an efficient register allocation mechanism, reduce the percentage of hotspots in the register file and uniformly distribute the heat. As a result, the thermal profile and reliability of the device is clearly improved. Simulation results show that the proposed flow achieved 91% reduction of hotspots and 11% reduction of the peak temperature.

- [6] “*Performance and Energy Trade-offs Analysis of L2 on-Chip Cache Architectures for Embedded MPSoCs*”. Aly, Mohamed M. Sabry ; Ruggiero, Martino ; Garcia del Valle, Pablo.

Proceedings of the 20th symposium on Great lakes symposium on VLSI, 2010, p. 305–310, ISBN: 978-1-4503-0012-4. Providence, Rhode Island, USA. May 16th-18th 2010.

On-chip memory organization is one of the most important aspects that can influence the overall system behavior in multi-processor systems. Following the trend set by high-performance processors, high-end embedded cores are moving from single-level on chip caches to a two-level on-chip cache hierarchy. Whereas in the embedded world there is general consensus on L1 private caches, for L2 there is still not a dominant architectural paradigm. Cache architectures that work for high performance computers turn out to be inefficient for embedded systems (mainly due to power-efficiency issues). This paper presents a virtual platform for design space exploration of L2 cache architectures in low-power Multi-Processor-Systems-on-Chip (MPSoCs). The tool contains several L2 caches templates, and new architectures can be easily added using our flexible plugin system. Given a set of constraints for a specific system (power, area, performance), our tool will perform extensive exploration to find the cache organization that best suits our needs. Through some practical experiments, we show how it is possible to select the optimal L2 cache, and how this kind of tool can help designers avoid some common misconceptions. Benchmarking results in the experiments section will show that for a case study with multiple processors running communicating tasks allocated on different cores, the private L2 cache organization still performs better than the shared one.

- [7] “3D-ICE: Fast Compact Transient Thermal Modeling for 3D-ICs with Inter-Tier Liquid Cooling”. Sridhar, Mahankali ; Raj, Arvind ; Vincenzi, Alessandro ; Ruggiero, Martino ; Brunswiler, Thomas ; Atienza Alonso, David. Proceedings of the 2010 International Conference on Computer-Aided Design (ICCAD 2010), vol. 1, num. 1, 2010, p. 1-8. New York: ACM and IEEE Press, 2010. San Jose, California, USA. November 7th-11th, 2010.

Three dimensional stacked integrated circuits (3D ICs) are extremely attractive for overcoming the barriers in interconnect scaling, offering an opportunity to continue the CMOS performance trends for the next decade. However, from a thermal perspective, vertical integration of high-performance ICs in the form of 3D stacks is highly demanding since the effective areal heat dissipation increases with number of dies (with hotspot heat fluxes up to $250\text{W}/\text{cm}^2$) generating high chip temperatures. In this context, inter-tier integrated microchannel cooling is a promising and scalable solution for high heat flux removal. A robust design of a 3D IC and its subsequent thermal management depend heavily upon accurate modeling of the effects of liquid cooling on the thermal behavior of the IC during the early stages of design. In this paper we present 3D-ICE, a compact transient thermal model (CTTM) for the thermal simulation of 3D ICs with multiple inter-tier microchannel liquid cooling. The proposed model is compatible with existing thermal CAD tools for ICs, and offers significant speed-up (up to $\times 975$) over a typical commercial computational fluid dynamics simulation tool while preserving accuracy (i.e., maximum temperature error of 3.4%). In addition, a thermal simulator has been built based on 3D-ICE, which is capable of running in parallel on multicore architectures, offering further savings in simulation time and demonstrating efficient parallelization of the proposed approach.

- [8] “Compact Transient Thermal Model for 3D ICs with Liquid Cooling via Enhanced Heat Transfer Cavity Geometries”. Sridhar, Mahankali ; Raj, Arvind ; Vincenzi, Alessandro ; Ruggiero, Martino ; Brunswiler, Thomas ; Atienza Alonso, David. Proceedings of the 16th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC’10), vol. 1, num. 1, 2010, p. pp. 5 – 11. New York: IEEE Press, 2010. Barcelona, Spain. October 6th-8th, 2010.

The advent of 3D stacked ICs with accumulating heat fluxes stresses thermal reliability and is responsible for temperature driven performance deterioration of the electronic systems Hot spots with power densities typically rising up to $250\text{W}/\text{cm}^2$ are not acceptable, with the result of limited per-

formance improvement in next generation high performance microprocessor stacks. Unfortunately traditional back-side cooling only scales with the chip stack footprint, but not with the number of tiers. Direct heat removal from the IC dies via inter-tier liquid cooling is a promising solution to address this problem. In this regard, a thermal-aware design of a 3D IC with liquid cooling for optimal electronic performance and reliability requires fast modeling and simulation of the liquid cooling during the early stages of the design. In this paper, we propose a novel compact transient thermal modeling (CTTM) scheme for liquid cooling in 3D ICs via microchannels and enhanced heat transfer cavity geometries such as pin-fin structures. The model is compatible with the existing thermal-CAD tools for ICs and offers significant speed-up over commercial computational fluid dynamics simulators ($\times 13478$ for pin-fin geometry with 1.1% error in temperature). In addition, the model is highly flexible and it provides a generic framework in which heat transfer coefficient data from numerical simulations or existing correlations can be incorporated depending upon the speed/accuracy needs of the designer. We have also studied the effects of using different techniques for the estimation of heat transfer coefficients on the accuracy of the model. This study highlights the need to consider developing flow conditions to accurately model the temperature field in the chip stack. The use of correlation data from fully developed flows only results in temperature error as high as 9 K (about 41%) near the inlet.

- [9] “Fuzzy Control for Enforcing Energy Efficiency in High-Performance 3D Systems”. Sabry Aly, Mohamed Mostafa, Ayse Kivilcim Coskun, David Atienza Alonso. Proceedings of the 2010 International Conference on Computer-Aided Design (ICCAD 2010), vol. 1, num. 1, 2010, p. 10-16. New York: IEEE Press, 2010. San Jose, California, USA. November 7th-11th, 2010.

3D stacked circuits reduce communication delay in multicore system-on-chips (SoCs) and enable heterogeneous integration of cores, memories, sensors, and RF devices. However, vertical integration of layers exacerbates the reliability and thermal problems, and cooling is a limiting factor in multitier systems. Liquid cooling is a highly efficient solution to overcome the accelerated thermal problems in 3D architectures; however, liquid cooling brings new challenges in modeling and runtime management. This paper proposes a novel controller for improving energy efficiency and reliability in 3D systems through liquid cooling management and dynamic voltage frequency scaling (DVFS). The proposed fuzzy controller adjusts the liquid flow rate at runtime to match the cooling demand for preventing energy wastage of over-cooling and for maintaining a stable thermal profile. The DVFS decisions provide chip-level energy savings and help balancing the temperature across the system. Experimental results on 8- and 16-core multicore SoCs show that the controller prevents the system to exceed the given threshold temperature while reducing cooling energy by up to 50% and system-level energy by up to 21% in comparison to using a static worst-case flow rate setting.

- [10] “Thermal-Aware Compilation for Register Window-Based Embedded Processors”. Aly, Sabry ; Mostafa, Mohamed ; Rodrigo, Ayala ; Luis, José ; Atienza Alonso, David. IEEE Embedded Systems Letters, vol. 2, num. 4, 2010, p. 213-217, Institute of Electrical and Electronics Engineers, 2010. ISSN: 1943-0663.

The development of compiler-based mechanisms to optimize the thermal profile of large register files to improve the processor reliability has become an important issue. Thermal hotspots have been known to cause severe reliability issues, while the thermal profile of the devices is also related to the leakage power consumption and the cooling cost. Register window-based architectures provide a relatively large register files. However, such large register files are not designed or utilized for thermal balancing or reliability enhancement. In this paper, we propose a compilation flow that utilizes the register windows to reduce optimize the thermal profile and to reduce the hotspots. As a result, the thermal profile and reliability of the device is clearly improved. Simulation results show that the

proposed flow achieves up to 91% reduction of hotspots and 11% reduction of the peak temperature in embedded processors.

- [11] “Evaluating OpenMP Support Costs on MPSoCs”. Marongiu A., Burgio P., Benini L.; Proceedings of the 13th Euromicro Conference on Digital Systems Design (DSD 2010). Lille, France. September 1st-3rd, 2010.

The ever-increasing complexity of MPSoCs is making the production of software the critical path in embedded system development. Several programming models and tools have been proposed in the recent past that aim at facilitating application development for embedded MPSoCs. OpenMP is a mature and easy-to-use standard for shared memory programming, which has recently been successfully adopted in embedded MPSoC programming as well. To achieve performance, however, it is necessary that the implementation of OpenMP constructs efficiently exploits the many peculiarities of MPSoC hardware. In this paper we present an extensive evaluation of the cost associated with supporting OpenMP on such a machine, investigating several implementative variants that efficiently exploit the memory hierarchy. Experimental results on different benchmarks confirm the effectiveness of the optimizations in terms of performance improvements.

- [12] “Vertical Stealing: Robust, Locality-Aware Do-All Workload Distribution for 3D MPSoCs”; A. Marongiu, P. Burgio, L. Benini. Proceedings of the 2010 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems. October 24th-28th, 2010. Scottsdale, Arizona, USA.

In this paper we address the issue of efficient doall workload distribution on a embedded 3D MPSoC. 3D stacking technology enables low latency and high bandwidth access to multiple, large memory banks in close spatial proximity. In our implementation one silicon layer contains multiple processors, whereas one or more DRAM layers on top host a NUMA memory subsystem. To obtain high locality and balanced workload we consider a two-step approach. First, a compiler pass analyzes memory references in a loop and schedules each iteration to the processor owning the most frequently accessed data. Second, if locality-aware loop parallelization has generated unbalanced workload we allow idle processors to execute part of the remaining work from neighbors by implementing runtime support for work stealing.

- [13] “Exploring Programming Model-Driven QoS Support for NoC-based Platforms”. Joven J., Marongiu A., Angiolini F., Benini L., De Micheli G.; Proceedings of the 2010 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS 2010). Scottsdale, Arizona, USA. October 24th-29th, 2010.

Networks-on-Chip (NoCs) are being increasingly considered as a central enabling technology to communication-centric designs as more and more IP blocks are integrated on the same SoC. Embedded applications, in turn, are becoming extremely sophisticated, and often require guaranteed levels of service and performance. The complex and non-uniform nature of network traffic generated by parallel applications running on a large number of possibly heterogeneous IPs makes a strong case for providing Quality of Service (QoS) support for traffic streams over the NoC infrastructure. In this paper we consider an integrated hardware/software approach for delivering QoS at the application level. We designed NoC hardware support, low-level middleware and APIs which enable QoS control at the application level. Furthermore, we identify a set of programming abstractions useful to associate the notion of priority to each running task in the system. An initial implementation of this programming model is also presented, which leverages a set of extensions to a MPSoC-specific OpenMP compiler and runtime environment.

- [14] “An OpenMP Compiler for Efficient Use of Distributed Scratchpad Memory in MPSoCs”. Marongiu A., Benini L.; IEEE Transactions on Computers, Vol. 59, Issue 12. December 2010. DOI: 10.1109/TC.2010.199.

Most of today’s state-of-the-art processors for mobile and embedded systems feature on-chip scratchpad memories. To efficiently exploit the advantages of low-latency high-bandwidth memory modules in the hierarchy there is the need for programming models and/or language features that expose such architectural details. On the other hand, effectively exploiting the limited on-chip memory space requires the programmer to devise an efficient partitioning and distributed placement of shared data at the application level. In this paper we propose a programming framework that combines the ease of use of OpenMP with simple yet powerful language extensions to trigger array data partitioning. Our compiler exploits profiled information on array access count to automatically generate data allocation schemes optimized for locality of references.

- [15] “Performance and Energy Trade-offs Analysis of L2 on-Chip Cache Architectures for Embedded MPSoCs”. M. M. S. Aly, M. Ruggiero and P. Garcia del Valle. Proceedings of the 20th symposium on Great lakes symposium on VLSI, Providence, Rhode Island, USA. May 16th-18th, 2010

On-chip memory organization is one of the most important aspects that can influence the overall system behavior in multi-processor systems. Following the trend set by high-performance processors, high-end embedded cores are moving from single-level on chip caches to a two-level on-chip cache hierarchy. Whereas in the embedded world there is general consensus on L1 private caches, for L2 there is still not a dominant architectural paradigm. Cache architectures that work for high performance computers turn out to be inefficient for embedded systems (mainly due to power-efficiency issues). This paper presents a virtual platform for design space exploration of L2 cache architectures in low-power Multi-Processor-Systems-on-Chip (MPSoCs). The tool contains several L2 caches templates, and new architectures can be easily added using our flexible plug-in system. Given a set of constraints for a specific system (power, area, performance), our tool will perform extensive exploration to find the cache organization that best suits our needs. Through some practical experiments, we show how it is possible to select the optimal L2 cache, and how this kind of tool can help designers avoid some common misconceptions. Benchmarking results in the experiments section will show that for a case study with multiple processors running communicating tasks allocated on different cores, the private L2 cache organization still performs better than the shared one.

- [16] “Scalable Instruction Set Simulator for Thousand-Core Architectures Running on GPGPUs”. Shivani Raghav, Martino Ruggiero, David Atienza, Christian Pinto, Andrea Marongiu, Luca Benini. Proc. of the 2010 International Conference on High Performance Computing & Simulation (HPCS 2010), pages 459–466. June 28th–July 2nd, 2010. Caen, France.

Simulators are still the primary tools for development and performance evaluation of applications running on massively parallel architectures. However, current virtual platforms are not able to tackle the complexity issues introduced by 1000-core future scenarios. We present a fast and accurate simulation framework targeting extremely large parallel systems by specifically taking advantage of the inherent potential processing parallelism available in modern GPGPUs.

- [17] “Programming Model and Compiler Support for Efficient Data Distribution in PGAS MP-SoCs”. Andrea Marongiu & Luca Benini. Poster presented at Intel European Research and Innovation Conference (ERIC). September 21th–22nd, 2010. Braunschweig, Germany. <http://www.intel.com/corporate/education/emea/event/irc/>. See Fig. 1 page 19.

- [18] “Efficient Execution of Kahn Process Networks on CELL BE”. Iuliana Bacivarov, Wolfgang Haid, Kai Huang, Lars Schor & Lothar Thiele. Presented at the Summer School on Models for Embedded Signal Processing Systems. Lorentz Center, Leiden, The Netherlands. August 30th-September 3rd, 2010. <http://www.lorentzcenter.nl/lc/web/2010/427/presentations/Iuliana-cell.pdf>.
- [19] “Distributed Operation Layer... A Practical Perspective”. Iuliana Bacivarov, Wolfgang Haid, Kai Huang & Lothar Thiele. Presented at the Summer School on Models for Embedded Signal Processing Systems. Lorentz Center, Leiden, The Netherlands. August 30th-September 3rd, 2010. <http://www.lorentzcenter.nl/lc/web/2010/427/presentations/Iuliana-demo.pdf>.

Scientific Contributions Published During the 2nd Reporting Period (2011)

- [20] “Thermally Optimal Stop-Go Scheduling of Task Graphs with Real-Time Constraints”. Pratyush Kumar, Lothar Thiele. Asia and South Pacific Design Automation Conference. Jan. 25th-28th 2011, Yokohama, Japan.

Dynamic thermal management (DTM) techniques to manage the load on a system to avoid thermal hazards are soon becoming mainstream in today's systems. Several research works have studied the use of dynamic voltage scaling (DVS) to serve a set of independent real-time tasks with temperature/performance constraints. In this paper, we study another important class of DTM techniques: stop-go scheduling, to minimize peak temperature when scheduling an application modeled as a task-graph within a given makespan constraint. For a given static-ordering of execution of tasks, we derive the optimal schedule referred to as a JUST policy. We also prove that for periodic task-graphs the optimal temperature is independent of the chosen static-ordering. Simulation experiments validate the theoretical results.

- [21] “Worst-Case Temperature Analysis for Real-Time Systems”. Lothar Thiele, Iuliana Bacivarov, Jian-Jia Chen, Devendra Rai, Hoeseok Yang. Proc. of Design, Automation and Test in Europe (DATE '11), ACM and IEEE Press. Grenoble, France. March 14th-18th, 2011.

With the evolution of today's semiconductor technology, chip temperature increases rapidly mainly due to the growth in power density. For modern embedded real-time systems, it is crucial to estimate maximal temperatures in order to take mapping or other design decisions to avoid burnout, and still be able to guarantee meeting real-time constraints. This paper provides answers to the question: When work-conserving scheduling algorithms, such as earliest-deadline first (EDF), rate-monotonic (RM), deadline-monotonic (DM), are applied, what is the worst-case peak temperature of a real-time embedded system under all possible scenarios of task executions? We propose an analytic framework, which considers a general event model based on network and real-time calculus. This analysis framework has the capability to handle a broad range of uncertainties in terms of task execution times, task invocation periods, and jitter in task arrivals. Simulations show that our framework is a cornerstone to design real-time systems that have guarantees on both schedulability and maximal temperatures.

- [22] “Towards Thermally-Aware Design of 3D MPSoCs with Inter-Tier Cooling”, Mohamed M. Sabry, Arvind Sridhar, David Atienza, Yuksel Temiz, Yusuf Leblebici, Sylwia Szczukiewicz, Navid Borhani, John R. Thome, Thomas Brunschwiler, and Bruno Michel. Proc. of Design, Automation and Test in Europe (DATE '11), ACM and IEEE Press. Grenoble, France. March 14th-18th, 2011.

New tendencies envisage 3D Multi-Processor System-On-Chip (MPSoC) design as a promising solution to keep increasing the performance of the next-generation highperformance computing (HPC) systems. However, as the power density of HPC systems increases with the arrival of 3D MPSoCs, supplying electrical power to the computing equipment and constantly removing the generated heat is rapidly becoming the dominant cost in any HPC facility. Thus, both power and thermal/cooling implications play a major role in the design of new HPC systems, given the energy constraints in our society. Therefore, in this work we propose a new holistic thermally-aware design. This paper presents the exploration of novel cooling technologies, as well as suitable thermal modeling and system-level design methods, which are all necessary to develop 3D MPSoCs with inter-tier liquid cooling systems. As a result, we develop energy-efficient run-time thermal control strategies to achieve energy-efficient cooling mechanisms to compress almost 1 Tera nano-sized functional units into one cubic centimeter with a 10 to 100 fold higher connectivity than otherwise possible. The proposed thermally-aware design paradigm includes exploring the synergies of hardware-, software- and mechanical-based thermal control techniques as a fundamental step to design 3D MPSoCs for HPC systems. Our management strategy prevents the system from surpassing the given threshold temperature while achieving up to 67% reduction in cooling energy and up to 30% reduction in system-level energy in comparison to setting the flow rate at the maximum value to handle the worst-case temperature.

- [23] “3D Thermal-Aware Floorplanner for Many-Core Single-Chip Systems”, David Cuesta, Jose L. Risco, Jose L. Ayala, David Atienza, Proc. of 12th IEEE Latin-American Test Workshop (LATW11), IEEE Press. Beach of Porto de Galinhas, Brazil. March 27th-30th, 2011.

Heat removal and power density distribution delivery have become two major reliability concerns in 3D stacked technology. In this paper, we propose a thermal-driven 3D floorplanner. Our contributions include: (1) a novel multi-objective formulation to consider the thermal and performance constraints in the optimization approach; (2) an efficient Mixed Integer Linear Programming (MILP) representation of the floorplanning model; and (3) a smooth integration of the MILP model with an accurate thermal modelling of the architecture. The experimental work is conducted for two realistic many-core single-chip architectures: an homogeneous system resembling Intel’s SCC, and an improved heterogeneous setup. The results show promising improvements of the mean, peak temperature and the thermal gradient, with a reduced overhead in the wire length of the system.

- [24] “Thermal-Aware System Analysis and Software Synthesis for Embedded Multi-Processors”, Lothar Thiele, Lars Schor, Hoeseok Yang and Iuliana Bacivarov. Design Automation & Test in Europe. March 14th-18th, 2011. Grenoble, France. <ftp://ftp.tik.ee.ethz.ch/pub/people/lrschor/tsyb2011a.pdf>.

Nowadays, the reliability and performance of modern embedded multi-processor systems is threaten by the ever-increasing power densities in integrated circuits, and a new additional goal of software synthesis is to reduce the peak temperature of the system. However, in order to perform thermal-aware mapping optimization, the timing and thermal characteristics of every candidate mapping have to be analyzed. While the task of analyzing timing characteristics of design alternatives has been extensively investigated in recent years, there is still a lack of methods for accurate and fast thermal analysis. In order to obtain desired evaluation times, the system has to be simulated at a high abstraction level. This often results in a loss of accuracy, mainly due to missing knowledge of system’s characteristics. This paper addresses this challenge and presents methods to automatically calibrate high-level thermal evaluation methods. Furthermore, the viability of the methods for automated model calibration is illustrated by means of a novel high-level thermal evaluation method.

- [25] “*Distributed Application Layer: Towards Efficient and Reliable Programming of Many-Tile Architectures*”. Iuliana Bacivarov, Hoeseok Yang, Lars Schor, Devendra Rai, Sudhanshu Jha, and Lothar Thiele. Design Automation & Test in Europe, March 14th-18th, 2011. Grenoble, France. <http://conferenze.dei.polimi.it/depcp/proceedings/pdfs/3.pdf>.

Many-tile architectures are large lattices that have multi-processor system-on-chips (MPSoC) as components, and as such, they push the MPSoC concept one step further in terms of raw computational power and homogeneous/ heterogeneous application parallelism. This advanced degree of computational flexibility comes at expenses of new requirements in terms of programming paradigms. The distributed operation layer (DOL) (<http://www.tik.ee.ethz.ch/shapes/>) was successfully utilized for programming and optimizing the static mapping of one parallel application on different multi-processor platforms like PC clusters, MPARM, IBM Cell, and Atmel D940. However, to cope with multiple, dynamic, and concurrent applications and to optimize their execution while still guaranteeing real-time constraints, a new programming model, named distributed application layer (DAL) is being developed on top of DOL. Besides offering an efficient programming model for many-tile architectures, DAL provides faster application development and reliability awareness. This poster introduces all basic concepts of DAL, presenting the programming model, design principles, and a proof-of-concept implementation.

- [26] “*Towards Thermally-Aware Design of 3D MPSoCs with Inter-Tier Cooling*”. Mohamed Sabry, Arvind Sridhar, David Atienza Alonso and Yuksel Temiz. Design Automation & Test in Europe, March 14th-18th, 2011. Grenoble, France.

New tendencies envisage 3D Multi-Processor System-On-Chip (MPSoC) design as a promising solution to keep increasing the performance of the next-generation highperformance computing (HPC) systems. However, as the power density of HPC systems increases with the arrival of 3D MPSoCs, supplying electrical power to the computing equipment and constantly removing the generated heat is rapidly becoming the dominant cost in any HPC facility. Thus, both power and thermal/cooling implications play a major role in the design of new HPC systems, given the energy constraints in our society. Therefore, EPFL, IBM and ETHZ have been working within the CMOSAIIC Nano-Tera.ch program project in the last three years on the development of a holistic thermally-aware design. This paper presents the exploration in CMOSAIIC of novel cooling technologies, as well as suitable thermal modeling and system-level design methods, which are all necessary to develop 3D MPSoCs with inter-tier liquid cooling systems. As a result, we develop energy-efficient run-time thermal control strategies to achieve energy-efficient cooling mechanisms to compress almost 1 Tera nano sized functional units into one cubic centimeter with a 10 to 100 fold higher connectivity than otherwise possible. The proposed thermally-aware design paradigm includes exploring the synergies of hardware-, software- and mechanical-based thermal control techniques as a fundamental step to design 3D MPSoCs for HPC systems. More precisely, we target the use of inter-tier coolants ranging from liquid water and twophase refrigerants to novel engineered environmentally friendly nano-fluids, as well as using specifically designed micro-channel arrangements, in combination with the use of dynamic thermal management at system-level to tune the flow rate of the coolant in each micro-channel to achieve thermally-balanced 3D-ICs. Our management strategy prevents the system from surpassing the given threshold temperature while achieving up to 67% reduction in cooling energy and up to 30% reduction in system-level energy in comparison to setting the flow rate at the maximum value to handle the worst-case temperature.

- [27] “*Rigorous Component-Based System Design Using the BIP Framework*”. Ananda Basu, Saddek Bensalem, Marius Bozga, Jacques Combaz, Mohamad Jaber, Thanh-Hung Nguyen,

Joseph Sifakis. IEEE Software, vol. 28, no. 3, May/June 2011. DOI: <http://dx.doi.org/10.1109/MS.2011.27>.

Rigorous system design requires the use of a single powerful component framework allowing the representation of the designed system at different levels of detail, from application software to its implementation. The use of a single framework allows to maintain the overall coherency and correctness by comparing different architectural solutions and their properties. In this paper, we present the BIP (Behavior, Interaction, Priority) component framework which encompasses an expressive notion of composition for heterogeneous components by combining interactions and priorities. This allows description at different levels of abstraction from application software to mixed hardware/software systems. Then, we introduce a rigorous design flow that uses BIP as a unifying semantic model to derive from an application software, a model of the target architecture and a mapping, a correct implementation. Correctness of implementation is ensured by application of source-to-source transformations in BIP which preserve correctness of essential design properties. The design is fully automated and supported by a toolset including a compiler, the D-Finder verification tool and model transformers. We illustrate the use of BIP as a modeling formalism as well as crucial aspects of the design flow for ensuring correctness, through an autonomous robot case study.

- [28] “*Thermal-Aware System-Level Modeling and Management for Multi-Processor Systems-on-Chip*”. Francesco Zanini, David Atienza, Luca Benini and Giovanni de Micheli. Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS’11), pages 2481–2484. Rio de Janeiro, Brazil, May 15th–18th, 2011. ISBN: 978-1-4244-9472-9/11. <http://infoscience.epfl.ch/record/162539>.

Multi-Processor Systems-on-Chip (MPSoCs) are penetrating the electronics market as a powerful, yet commercially viable, solution to answer the strong and steadily growing demand for scalable and high performance systems, at limited design complexity. However, it is critical to develop dedicated system-level design methodologies for multi-core architectures that seamlessly address their thermal modeling, analysis and management. In this work, we first formulate the problem of system-level thermal modeling and link it to produce a global thermal management formulation as a discrete-time optimal control problem, which can be solved using finite-horizon model-predictive control (MPC) techniques, while adapting to the actual time-varying unbalanced MPSoC workload requirements. Finally, we compare the system-level MPC-based thermal modeling and management approaches on an industrial 8-core MPSoC design and show their different trade-offs regarding performance while respecting operating temperature bounds.

- [29] “*Thermal Analysis and Active Cooling Management for 3D MPSoCs*”. M. Sabry, D. Atienza and A. K. Coskun. Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS’11), pages 2237–2240. Rio de Janeiro, Brazil, May 15th–18th, 2011. ISBN: 978-1-4244-9472-9/11. <http://infoscience.epfl.ch/record/163372>.

3D stacked architectures reduce communication delay in multiprocessor system-on-chips (MPSoCs) and allowing more functionality per unit area. However, vertical integration of layers exacerbates the reliability and thermal problems, and cooling is a limiting factor in multi-tier systems. Liquid cooling is a highly efficient solution to overcome the accelerated thermal problems in 3D architectures. However, liquid cooling brings new challenges in modeling and run-time management. This paper proposes a design-time/run-time thermal management policy for 3D MPSoCs with inter-tier liquid cooling. First, we perform a design-time analysis to estimate the thermal impact of liquid cooling and dynamic voltage frequency scaling (DVFS) on 3D MPSoCs. Based on this analysis, we define a set of management rules for run-time thermal management. We utilize these rules to control and adjust the liquid flow rate in order to match the cooling demand for preventing energy wastage of

overcooling, while maintaining a stable thermal profile in the 3D MPSoCs. Experimental results on multi-tier 3D MPSoCs show that proposed design-time/run-time management policy prevents the system to exceed the given threshold temperature while reducing cooling energy by 50% on average and system-level energy by 18% on average in comparison to using a static worstcase flow rate setting.

- [30] “GPGPU-Accelerated Parallel and Fast Simulation of Thousand-core Platforms”, Christian Pinto, Shivani Raghav, Andrea Marongiu, Martino Ruggiero, David Atienza, Luca Benini. IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing 2011 (CCGRID2011). Newport Beach, California. May 23rd-26th 2011. DOI: 10.1109/CC-Grid.2011.64. <http://infoscience.epfl.ch/record/164471/files/ccgrid11.pdf>.

The multicore revolution and the ever-increasing complexity of computing systems is dramatically changing system design, analysis and programming of computing platforms. Future architectures will feature hundreds to thousands of simple processors and on-chip memories connected through a Network-on-Chip (NoC). Architectural simulators will remain primary tools for design space exploration, software development and performance evaluation of these massively parallel architectures. However, architectural simulation performance is a serious concern, as virtual platforms and simulation technology are not able to tackle the complexity of thousands of core future scenarios. The main contribution of this paper is the development of a new simulation approach and technology for many core processors which exploit the enormous parallel processing capability of low-cost and widely available General Purpose Graphic Processing Units (GPGPU). The simulation of many-core architectures exhibits indeed a high level of parallelism and is inherently parallelizable, but GPGPU acceleration of architectural simulation requires an in-depth revision of the data structures and functional partitioning traditionally used in parallel simulation. We demonstrate our GPGPU simulator on a target architecture composed by several cores (i.e. ARM ISA based), with instruction and data caches, connected through a NoC. Our experiments confirm the feasibility of our approach.

- [31] “Die-Level TSV Fabrication Platform for CMOS-MEMS Integration”. Yuksel Temiz, Michail Zervas, Carlotta Guiducci, and Yusuf Leblebici. Proceedings of the 16th International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers 2011). June 5th-9th, 2011. Beijing, China.

This paper reports a new post-CMOS processing platform for die-level through-silicon-via (TSV) fabrication, based on wafer reconstitution from embedded dies, parylene deposition, stencil lithography, and bottom-up electroplating. The goal of this work is to develop a heterogeneous 3D-integration technique for the applications requiring CMOS-MEMS integration with vertical interconnections.

- [32] “Cool shapers: Shaping Real-Time Tasks for Improved Thermal Guarantees”, Pratyush Kumar and Lothar Thiele, Design Automation Conference 2011, June 5th-10th, 2011. San Diego, California, USA. <ftp://blackbox.ethz.ch/pub/people/kumarpr/DAC11.pdf>

With increasing power densities, managing on-chip temperatures has become an important design challenge. We propose a novel approach to this problem with the use of shapers to dynamically and selectively insert idle times during the execution of hard real-time jobs on a single speed processor. For the class of leaky bucket shapers which have a light-weight implementation, we derive the shaper such that no job misses its real-time deadline and the peak temperature is optimally reduced. The analysis and design of such shapers allows for dynamically variable streams of jobs; for instance, periodic streams with jitter. We extend our results to consider non-zero power and timing overhead in transitioning to the idle mode. With experimental results, we demonstrate that the proposed approach provides a large improvement: on average 8K peak temperature reduction or 40% increase in utilization for a given peak temperature.

- [33] “*Temperature Predictability in Multi-Core Real-Time Systems*”. Iuliana Bacivarov, Lars Schor, Hoeseok Yang, Devendra Rai, and Lothar Thiele. Presentation at the DAC Workshop on Multiprocessor System on Chip for Cyber Physical Systems: Programmability, Run-Time Support and Hardware Platforms for High Performance Embedded Applications. At DAC 2011, San Diego, California, USA. June 9th, 2011. http://www.dac.com/workshops+_+colocated+events.aspx?event=104&topic=15.
- [34] “*Thermal-Aware Design of Real-Time Multi-Core Embedded Systems*”. Iuliana Bacivarov, Lars Schor, Hoeseok Yang, Devendra Rai, and Lothar Thiele. Presentation at the Artemis Workshop on Mapping of Applications to MPSoCs (Map2MPSoC 2011). Rheinfels Castle, St. Goar, Germany. June 28th-29th, 2011. <http://www.artist-embedded.org/artist/Program,2298.html>.
- [35] “*End-to-End Delay Minimization in Thermally Constrained Distributed Systems*”, Pratyush Kumar and Lothar Thiele, Euromicro Conference on Real-Time Systems (ECRTS 2011). July 6th–8th, 2011. Porto, Portugal.

With ever-increasing power densities, managing on-chip temperatures by optimizing mapping and scheduling of tasks is becoming increasingly necessary. We study the minimization of end-to-end delay for thermally constrained scheduling of an application that is specified as a task graph and is executing on parallel processors without speed scaling. We show that task graph scheduling on thermally constrained systems is monotonic, i.e., delaying the execution of a task longer than necessary cannot lead to the early completion of any other task. Using this monotonicity principle, we design the provably optimal schedule for a given mapping, called the JUST schedule. The JUST schedule can be easily implemented using temperature sensors. We then present different thermal-aware modifications to standard mapping heuristics and evaluate them on a large set of problem instances. The experimental results illustrate that with simple thermal-aware modifications, mappings with much smaller end-to-end delay can be identified.

- [36] “*Rigorous System Level Modeling and Analysis of Mixed HW/SW Systems*”. Paraskevas Bourgos, Ananda Basu, Marius Bozga, Saddek Bensalem, Kai Huang and Joseph Sifakis. ACM/IEEE Ninth International Conference on Formal Methods and Models for Codesign (MEMOCODE’11). July 11th-13th, 2011. Cambridge, United Kingdom.

A grand challenge in complex embedded systems design is developing methods and tools for modeling and analyzing the behavior of an application software running on multicore or distributed platforms. We propose a rigorous method and a tool chain that allows to obtain a faithful model representing the behavior of a mixed hardware/software system from a model of its application software and a model of its underlying hardware architecture. The system model can be simulated and analyzed for validation of both functional and extra-functional properties. The tool chain uses DOL (Distributed Operation Layer, from ETH Zürich) as the frontend for specifying the application software and hardware architecture, and BIP (Behavior Interaction Priority, from VERIMAG) as the modeling and analysis framework. It is illustrated through the construction of system models of MJPEG and MPEG2 decoder applications running on MARM, a multicore architecture from Università di Bologna.

- [37] “*Hierarchical Thermal Management Policy for High-Performance 3D Systems with Liquid Cooling*”. Francesco Zanini, Mohamed Sabry, David Atienza Alonso and Giovanni De Micheli. IEEE Journal of Emerging and Selected Topics in Circuits and Systems (JETCAS), volume 1, issue 1, July 2011. DOI: 10.1109/JETCAS.2011.2158272. <http://infoscience.epfl.ch/record/165809>.

3-Dimensional integrated circuits and systems are expected to be present in electronic products in the short term. We consider specifically 3-D multi-processor systems-onchip (MPSoCs), realized by stacking silicon CMOS chips and interconnecting them by means of through-silicon vias (TSVs). Because of the high power density of devices and interconnect in the 3D stack, thermal issues pose critical challenges, such as hot-spot avoidance and thermal gradient reduction. Thermal management is achieved by a combination of active control of on-chip switching rates as well as active interlayer cooling with pressurized fluids. In this paper, we propose a novel online thermal management policy for high-performance 3D systems with liquid cooling. Our proposed controller uses a hierarchical approach with a global controller regulating the active cooling and local controllers (on each layer) performing dynamic voltage and frequency scaling (DVFS) and interacting with the global controller. Then, the online control is achieved by policies that are computed off-line by solving an optimization problem that considers the thermal profile of 3D-MPSoCs, its evolution over time and current time-varying workload requirements. The proposed hierarchical scheme is scalable to complex (and heterogeneous) 3D chip stacks. We perform experiments on a 3D-MPSoC case study with different interlayer cooling structures, using benchmarks ranging from web-accessing to playing multimedia. Results show significant advantages in terms of energy savings that reaches values up to 50% versus state-of-the-art thermal control techniques for liquid cooling, and thermal balance with differences of less than 10°C per layer.

- [38] “Attaining Single-Chip, High-Performance Computing Through 3D Systems with Active Cooling”. Ayse K. Coskun, David Atienza, Mohamed Sabry and Jie Meng. IEEE Micro, vol. 16, num. 5, 2011, pages 25–35. <http://infoscience.epfl.ch/record/165525>.

Three-dimensional (3D) stacking is an attractive method for designing large manycore chips as it provides high transistor integration densities, improves manufacturing yield due to smaller chip area, reduces wirelength and capacitance, and enables heterogeneous integration of different technologies on the same chip. Stacking, however, significantly increases the thermal resistivity and the on-chip temperatures. In fact, temperature is among the major manufacturing challenges for 3D design. Active cooling, where the chip is cooled through the liquid flowing in built-in microchannels or through a cold plate, has emerged as a viable cooling alternative for high-performance 3D manycore systems. Liquid cooling is more efficient in removing the heat in comparison to conventional cooling methods with heat sinks and fans. Nevertheless, the dynamically changing nature of workloads running on manycore systems require runtime techniques to enable energy-efficient, reliable, and high-performance operation of liquid-cooled 3D systems. This article focuses on the benefits and the challenges of 3D design, and discusses novel techniques to integrate predictive cooling control with chip-level thermal management methods such as job scheduling and voltage frequency scaling. The key message is that 3D liquid-cooled systems with intelligent runtime management provide an energy-efficient solution to designing future single-chip high-performance manycore architectures.

- [39] “Real-Time Analysis of Servers for General Job Arrivals”. Pratyush Kumar, Jian-Jia Chen, Lothar Thiele, Andreas Schranzhofer and Giorgio C. Buttazzo. Proceedings of the 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2011), August 28th-31st, 2011. Toyama, Japan.

Several servers have been proposed to schedule streams of aperiodic jobs in the presence of other periodic tasks. Standard schedulability analysis has been extended to consider such servers. However, not much attention has been laid on computing the worst-case delay suffered by a given stream of jobs when scheduled via a server. Such analysis is essential for using servers to schedule hard real-time tasks. We illustrate, with examples, that well established resource models, such as supply bound function and models from Real-Time Calculus, do not tightly characterize servers. In this work, we analyze the server algorithm of the Constant Bandwidth Server and compute a provably tight

resource model of the server. The approach used enables us to differentiate between the soft and hard variants of the server. A similar approach can be used to characterize other servers; the final results for which are presented.

- [40] “System-Level Power and Timing Variability Characterization To Compute Thermal Guarantees”. Pratyush Kumar and Lothar Thiele, CODES/ISSS 2011 at Embedded Systems Week, 9th-14th October, 2011. Taipei, Taiwan.

With ever-increasing power densities, temperature management using software and hardware techniques has become a necessity in the design of modern electronic systems. Such techniques have to be validated and optimized with respect to the thermal guarantee they provide, i.e., a safe upper-bound on the peak temperature of the system under all operating conditions. The computation of such a guarantee depends on the power and timing characteristics of the system. In this paper, we present formalisms to capture such characteristics at the system-level and provide an analytical technique to compute a provably safe upper-bound on the peak temperature. The proposed characterization and analysis is general in that it considers an impulse-response-based thermal model, task-dependent power consumption, tasks with dynamic arrival patterns and variable resource demand, and a scheduling policy expressed as a hierarchical composition of several commonly used policies.

Programming Model and Compiler Support for Efficient Data Distribution in PGAS MPSoCs

Andrea Marongiu, Luca Benini
DEIS - University of Bologna
Italy



INTRODUCTION

MPSoCs are heterogeneous systems with a complex memory hierarchy, often organized as a distributed shared memory, where a Global Address Memory Space (GAS) is physically partitioned in several modules that are accessed at different costs from different cores (NUMA).

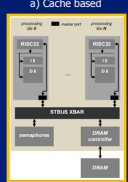
OpenMP is a mature standard for shared memory parallel programming, but it was intended for cache coherent SMP machines → it lacks any architectural awareness and constructs to control data placement and locality.

We propose an MPSoC-suitable implementation, with several custom extensions:

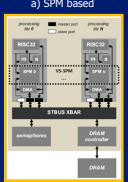
- Support for array data partitioning and distributed placement
- Streamlined embedded MPSoC suitable implementation of the necessary support for software address translation in absence of MMUs
- A compiler optimization pass for feedback-based array tile placement across multiple SPMs with NUMA organization
- Language features and run-time support to enable re-use of SPM space by means of DMA movements


a) Cache based



a) SPM based



PGAS




Our Target Architectural Templates

- Configurable number of processing tiles
- Partitioned Global Address Space (PGAS)
- Physically distributed shared memory
- Local scratchpads (SPM) to each core
- Globally non-coherent data caches
- SW. OS-less (custom support middleware)


OpenMP Extensions for array partitioning and data distribution

ARRAY PARTITIONING

```
#pragma omp distributed (A, tilesize)
int regular(A)[8];
#pragma omp distributed (regular,AsB/MPROCS)
int irregular(O);
#pragma omp distributed (irregular)
#pragma omp parallel for tiled(irregular) split(regular)
for (i=0; i<N; i++)
irregular[regular[i]] = ...;
```



cyclic distribution



block distribution

The split clause

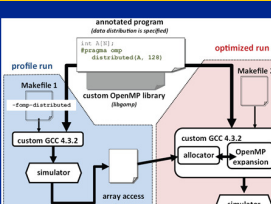
```
tileid[0][0] = 0x22000028;
tileid[0][1] = 0x22100028;
tileid[0][2] = 0x22200028;
tileid[0][3] = 0x22300028;
double a[8];
GOMP_get_split_tile_base_addr
(int dvar_id, int tile)
{
double * base;
base = tiles[dvar_id][tile];
return (double *)base;
}
main_omp_fn_0 (...)
{
double * base;
base = GOMP_get_split_tile_base_addr
(<distvar_id, <tile>);
for (i=0; i<2; i++) {
(*base)[i] = foo(i);
}
```

The tiled clause

```
tileid[0][0] = 0x22000028;
tileid[0][1] = 0x22100028;
tileid[0][2] = 0x22200028;
tileid[0][3] = 0x22300028;
double a[8];
GOMP_access_tiled_array
(int offset, int dvar_id)
{
double * base;
base = GOMP_access_tiled_array
(<offset, <distvar_id>);
tile = offset/tilesize[dvar_id];
base = tiles[dvar_id][tile];
return (double *) (base + offset);
}
```

OpenMP Extensions for DMA Movements

Profile-based data distribution



annotated program (data distribution is enabled) → profile run → Makefile-1 → custom OpenMP library (libgomp) → custom GCC 4.3.2 → simulator → array access trace → allocator → OpenMP extension → Makefile-2 → optimized run

The profile_locality clause

```
#pragma omp profiled_locality tiled(A) \
copyarrayin(A) copyarrayout(A)
for (i=0; i<N; i++)
#pragma omp parallel for
for (j=0; j<M; j++) { ... }
```

The copyarrayin and copyarrayout clauses

```
#pragma omp parallel for tiled(irregular) split(regular) \
copyarrayin(irregular) copyarrayout(irregular)
__builtin_gomp_copy_to_scratch
(uint src_addr, uint tsize);
__builtin_gomp_copy_to_extmem
(uint src_addr, uint tsize);
```

The copytilein and copytileout clauses

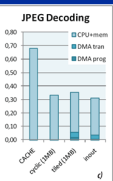
```
__builtin_gomp_single_copy_to_scratch
(uint src_addr, uint dst_addr, uint tsize);
__builtin_gomp_single_copy_to_extmem
(uint src_addr, uint dst_addr, uint tsize);
```

The inout scheduling clause

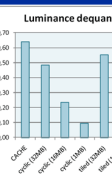
```
int regular(A)[8];
#pragma omp distributed(regular,B)
#pragma omp parallel for schedule(dynamic,1)
for (i=0; i<N; i++)
for (j=0; j<B; j++)
irregular[regular[i][j]] = ...;
#pragma omp for schedule(dynamic, 1, inout(regular))
```

EXPERIMENTAL RESULTS

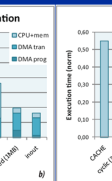
JPEG Decoding



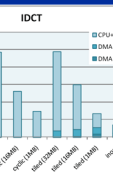
Luminance quantization



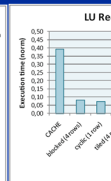
IDCT



LU Reduction



FFT



NCC

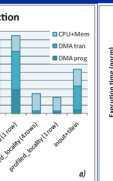


Figure 1: Poster Presented at Intel ERIC – Braunschweig, Germany – Sep. 19th-23rd, 2010