Thermal/Performance Trade-off in Network-on-Chip Architectures

Davide Zoni, Simone Corbetta and William Fornaciari

Politecnico di Milano – Dipartimento di Elettronica e Informazione

Via Ponzio 34/5, 20133 Milano, Italy

Email: {zoni,scorbetta,fornacia}@elet.polimi.it

Abstract—**Multi-core architectures are a promising paradigm to exploit the huge integration density reached by high-performance systems. Indeed, integration density and technology scaling are causing undesirable operating temperatures, having net impact on reduced reliability and increased cooling costs. Dynamic Thermal Management (DTM) approaches have been proposed in literature to control temperature profile at run-time, while design-time approaches generally provide floorplan-driven solutions to cope with temperature constraints. Nevertheless, a suitable approach to collect performance, thermal and reliability metrics has not been proposed, yet. This work presents a novel methodology to jointly optimize temperature/performance trade-off in reliable high-performance parallel architectures with security constraints achieved by workload physical isolation on each core. The proposed methodology is based on a linear formal model relating temperature and duty-cycle on one side, and performance and duty-cycle on the other side. Extensive experimental results on real-world use-case scenarios show the goodness of the proposed model, suitable for design-time system-wide optimization to be used in conjunction with DTM techniques.**

I. INTRODUCTION

Aggressive technology scaling has lead continuous miniaturization of transistors, making modern processors experiencing an exponential increase of performance in terms of clock rate, however with power consumption going as faster as clock rate [18]. Higher power consumption density in lower area regions makes operating temperatures increase up to the point reliability is mainly affected by thermal hot-spots: it has been shown that 50% of failures in CMOS integrated circuits are due to thermal issues [16], [22]. The transition to multi-core architectures introduced an opportunity for performance to grow faster than power consumption [5], allowing for a fine grain control on power densities and operating temperatures. Nevertheless, the increasing performance attained by multi-core and many-core processors are again raising the issue of integration capability and inter-core communication for future Multi-Processors System-on-Chip (MPSoC) design [12]. Network-on-Chip (NoC) architectures [3] have been proposed to cope with increasing performance requirements in massively parallel systems, but routers and link drivers consume a non-negligible amount of chip power [14], with a net impact on the chip temperature. In particular a few commercial designs show that the NoC can contribute up to 28% of total chip power [10]. Thermal Design Power (TDP) is the most challenging design constraint that accounts for and, sometimes, determines

the feasibility of the final system. In this perspective, thermal issues must be accounted at each design step, both at early design stages and at run-time. However, one of the main challenges in this perspective is to find a set of appropriate metrics that allows to manage and optimize such sensible chip design aspects, i.e. performance, thermal profile, power.

This work addresses thermal performance trade-off in a particular scenario, where multicore architectures are used to ensure security in critical web-service transactions and each workload/application must be mapped on a single core with no overlap. Traditionally, virtualization techniques are used to provide a logical separation between workloads on the same system, mainly for security reasons. However, such isolation is mild since it relies on virtual machine software components, that are usually designed for performance and can be violated quite easily [21]. Unlike existing software isolation techniques, the new trend on security seems to address isolation problem by a proper set of hardware modules, where workloads are physically isolated on a single core [2].

A. Novel contributions

The novel contributions of the work presented in this paper are many-fold. This research work focuses on the joint optimization of performance and temperature profile in multi-core architectures with NoC interconnect. The objective of our work is to analyse the thermal/performance trade-off in parallel architectures employed in high-performance server systems. To this extent, the following contributions are discussed in this paper:

- *Thermal/performance optimization* - an optimization methodology to jointly deal with performance and thermal profile trade-off is proposed as design-time optimization framework. The proposed work is general enough to be employed to constraint chip temperature, while maximizing core performance allowing for minimum core-to-core performance differences. In particular, we want to obtain a per-core maximum performance level with two conditions: chip temperature is maintained below a certain threshold and performance differences between cores are minimized;
- *Performance and duty-cycle* - we use clock-gating to control the performance of the cores; a valuable relationship between the applied clock-gating level (i.e., duty-cycle specification) and the performance degradation is then

proposed and validated against extensive experimental results;

- *Temperature and duty-cycle* - a valuable relationship between the applied clock-gating level and the operating temperature is proposed and validated against a rich set of experimental results;
- *System-wide optimization* - by employing the chip topology, we introduce the concept of topological rings to deal with thermal and performance trade-off. We propose a novel system-wide optimization methodology for multi-core architectures underpinned by the new concept of topological ring;
- *Real use case scenario* - to demonstrate the validity of the proposed approach, we cast our methodology on a specific available multi-core architecture [20]. After an in-depth use case analysis, we have exploited the multi-core architecture specificities providing strengthen results on our methodology.

To attain the contributions of the proposed research work, we propose two different tools. First, we developed a linear optimization model, to deal with thermal/performance trade-off. Moreover, we have cast both the performance and temperature empirical relations, extracted from data, as linear equations. Second, experimental results have been collected through an ad-hoc simulation framework, capable of cycle-accurate and thermal simulation of multi-core architectures with standard NoC interconnects.

B. Paper structure

This paper is organized as follows. Section II will give a brief overview of the state-of-the-art thermal optimization techniques, both at design-time and run-time. Section III introduces the proposed formal model for joint thermal/performance optimization under either absolute temperature constraints. Experimental results are discussed in Section IV, and conclusions will be drawn in Section V.

II. RELATED WORKS

The reliability dependence on increasing operating temperatures of microelectronics systems makes the control of the temperature profile of utmost importance in multi-core processors. Thermal management refers to a set of techniques and design choices that leads to the optimization of the temperature profile of a chip: hard-faults mechanisms such as electromigration and stress-migration are known to be exponentially related to operating temperature [27]. Optimization techniques can be employed either at design-time or at run-time. The former approaches have the advantage of finer-grain control (e.g., circuit-level techniques or microarchitecture-level techniques) at the cost of reduced flexibility and increased silicon area. The latter approaches, on the other hand, has greater flexibility, but generally require additional software complexity (e.g., additional data structures to hold temperature information on a per-core granularity) and might have non-negligible effects on performance, without the opportunity to trade performance and thermal off in an easy way.

A. Design-time thermal optimization

Design-time thermal management techniques can be conveniently organized in two broad classes [15]: microarchitecture-level techniques and floorplanning optimizations. At the microarchitecture level we can find several works for general purpose applications, ranging from techniques targeting processor cores only, or techniques for on-chip memory caches. In the first case the processors can be restructured according to a cluster-based architecture, or by duplicating portions of the processor that are known to be thermal hot-spots. Functional units are duplicated in [9], with increased hardware area and cost: these units are used alternatively to reduce the stress on each single unit (e.g., an ALU or register files). Similar work has been done in [25] in which the only register file has been duplicated, and activity migration is directed toward the spare unit under dynamic thermal constraints. Functional units can also be resized to accommodate a lower power density [23], but with a reduction of the clock frequency and negative impact on processor performance. Floorplan can also be conveniently designed to accommodate thermal hot-spots as done in [19].

Design-time tools are generally required to perform predictions on the benefits of the thermal management solution under investigation, such that to modify where appropriate the entire design. A few works have tried to integrate performance, power and thermal analysis in a single framework. The `Polaris` framework [26] allows to estimate power and area of NoC-based designs, but does not allow to provide detailed power consumption profile for the processors and memory hierarchy. The work in [11] proposes an integrated framework for power, area and thermal modeling for large-scale computing systems. In this work, application traces are emulated rather than collected from cycle-accurate simulation, thus without considering the real behavior of reference use-case scenarios. The authors in [4] propose an integrated approach based on Virtutech Simics functional simulator, employing power and thermal models from real hardware characterization. The advantage of this approach relies on the possibility to develop, analyze and tune different control algorithms for thermal and power management, based on high-level Matlab descriptions. However, the power and thermal models are bound to a particular architecture and floorplan (an Intel©Xeon X7350 system), and also the simulation is not cycle-accurate. These aspects make the approach in [4] unsuitable for accurate thermal evaluation of MPSoC architectures with NoC communication channel running different core configurations and floorplans.

B. Dynamic thermal management

In the context of Dynamic Thermal Management (DTM), several approaches have been presented in literature for the run-time optimization of thermal profile in single-chip multiprocessor architectures. The major concern in this kind of works is the lack of an appropriate metric specifying the impact of temperature-related decisions on the performance

degradation of the system (e.g., on the impact of CPI). Indeed, several authors provide a methodology based on simple temperature predictive control to avoid exceeding a predefined threshold value [28]. History-based approaches in this sense have been proposed in [30]. The performance impact of many DTM techniques for high-performance microprocessors has been extensively discussed in [6].

III. PROPOSED METHODOLOGY

This section details the four main aspects of the methodology proposed in this paper. Before presenting in details the formal model, it is worth giving some basic definitions that will be used throughout the entire section. The reference architecture is multi-core and composed of *tiles* placed in a 2D-mesh topology. Each tile is composed of a processor core, a router and a L2 cache bank; the router is used to interface to the distributed (shared) L2 cache. The 2D-mesh topology is logically composed of a set $R := \{1, 2, ..., n_R\}$ of n_R rows and a set $C := \{1, 2, ..., n_C\}$ of n_C columns. We also consider a set $D := \{1, 2, ..., n_D\}$ of duty-cycle islands. A duty-cycle island is composed by a set of tiles with a common clock rate. Each tile belongs to one and only one duty-cycle island.

The remainder of this section is organized as follows: at first, an optimization linear model to deal with the thermal/performance trade-off is sketched in Section III-A; such model is underpinned by two formal analytical relations on temperature and performance. Temperature and performance linear relations are discussed in details in Section III-B and Section III-C, respectively. Last, Section III-D details how the 2D-mesh topology has been exploited to support design time thermal performance analysis.

A. Thermal/performance linear model

We consider three sets of variables for the optimization linear model. For each tile $(i, j) \in R \times C$, the integer variable $p_{i,j} \in \{0\%, 1\%, ..., 100\%\}$ represents the performance degradation level of the tile with respect to the base-case where performance is 100%, and $t_{i,j} \geq 0$ defines its temperature. It is worth to notice that we measure core performance degradation level with respect to the maximum performance of the same core. We employ clock-gating to tune performance of each core such that the maximum performance is intended as duty-cycle equal to 1, without any clock-gating action. For each island $d \in D$, the integer variable r_d specifies its duty-cycle, i.e. the fraction of time the core in the tile is active, with respect to the time clock-gating stops its execution. We aim at maximizing the minimum performance for each tile, as specified by the following objective function, where the max-min formulation is satisfied by Equation 2.

$$max\ q \qquad\qquad\qquad\qquad\qquad (1)$$
$$q \leq p_{i,j} \qquad\qquad \forall\ (i, j) \in R \times C \qquad (2)$$

The first constraint to bind the frequency of each tile to its own duty-cycle island is as follows:

$$p_{i,j} = 1 - r_{f(i,j)} \qquad\qquad \forall\ (i, j) \in R \times C, \qquad (3)$$
$$f := (i, j) \to D, \qquad (4)$$

where f represents a mapping function between the Cartesian coordinates (i, j) of the tile in the 2D-mesh topology, and the duty-cycle island. The proposed methodology is biased toward this function, and further details will be given in Section III-D. Temperature-aware designs constraint the maximum operating temperature to a predefined threshold temperature T_{max}, determining the reliability of the processor chip. This constraint can be defined as a simple relation, as follows:

$$t_{i,j} \leq T_{max} \qquad \forall\ (i, j) \in C \times R \qquad (5)$$

The optimization model presented so far sets a threshold temperature to the chip (Equation 5), meanwhile maximizing the performance of the worst-case task (Equation 1 and Equation 2). The result of this joint optimization lies in fairness of performance degradation across tiles belonging to different duty-cycle islands.

B. Thermal linear model

The linear model presented in Section III-A allows to maximize performance, under a maximum operating temperature constraint. However, the intrinsic simplicity of the maximum temperature requirement lacks of a suitable formulation to be employed in the linear optimization model. This section details a derived linear thermal equation that is meant to be employed in the proposed linear optimization model; this model is derived from extensive and accurate simulation measurements using a cycle-accurate simulation of homogeneous architectures (refer to Table I for more details on this). The thermal model of each tile is defined as:

$$t_{i,j} := \sum_{d \in D} (\alpha_d \cdot r_d) \qquad \forall\ (i, j) \in R \times C, \qquad (6)$$

where the temperature $t_{i,j}$ of tile (i, j) is linearly dependent on the duty-cycle r_d of island $d \in D$, and weighted by an unknown coefficient α_d, to be determined. In order to characterize Equation 6, i.e. quantifying α_d coefficients, we use a least square approach, since regressors are supposed to be independent. This means that duty-cycle islands are decoupled each other, with the advantage of finer grain control, but at increased hardware cost (associated to the control circuitry). To characterize the model, we have extracted a rich set of per-tile temperature measurements, using different duty-cycle combinations, using the cycle-accurate simulation framework presented in Section IV. Experimental results have shown a strong linear relation between regressors and temperature, strengthen by an analysis of the R^2 fitting coefficient, that is very close to 1. Moreover, experimental data generate a very well conditioned matrix A with $cond(A) \leq 10$ in all of conducted experiments on both 16 and 36 cores.
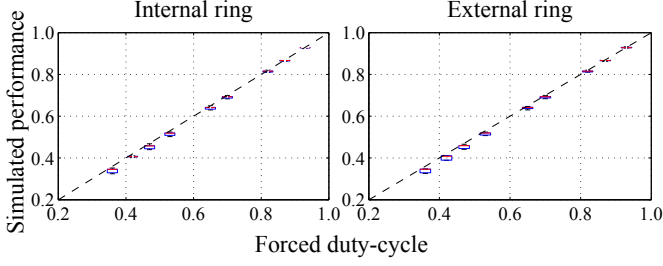
Fig. 1. Normalized simulated performance as a function of the applied clock-gating, for cores belonging to internal and external ring in 16-cores 2D-mesh.

Fig. 2. Rings in 16-cores and 36-cores 2D-mesh: a topological ring groups tiles that have similar thermal-related characteristics.

C. Performance linear model

The optimization model in Section III-A uses performance and temperature measurements to exploit the thermal/performance trade-off. This section details the linear relation that binds processor performance to the duty-cycle it belongs to. Notice that the validity of the proposed linear model is underpinned by the fact that the reference processor is an in-order core. Although the validity of the model is coupled with a specific and simple architecture, it is worth noticing that such in-order processors are still used in high-performance systems, such as Web servers or Data centers [20]. In addition, each core in the multi-core processor can be assumed to be isolated from the rest of the chip, because each core is assumed to serve a single request, to maximize response throughput and to ensure logical and physical security [2].

Processors run at a fixed clock frequency, and the performance is related to the number of committed instructions, bound to the level of duty-cycle specified by the island the processor belongs to. Moreover, for simple and only in-order cores without multi-thread capabilities, the clock rate is tightly coupled to all the executed instructions.

We have experimentally validated such relation by an extensive set of experiments on our cycle-accurate simulation framework using benchmarks from different test suites, finding a strong linear correlation between committed instructions and duty-cycle. Figure 1 shows the linear relationship between the number of committed instructions (*Simulated performance* on vertical axis) and the applied duty-cycle (*Forced clock-gating* on horizontal axis) for a 16-cores architecture in both internal and external topological rings. The dotted line represents the theoretical linear relation between committed instructions and clock-gating level, while the box-and-whiskers plots represent the simulated performance: for each clock-gating level, the maximum, minimum and median simulated performance are reported. The height of each box plot is tied to the variability of the simulated measurements: 50% of the simulated values fall in this interval. The width of the box plot, on the other hand, has no statistical meaning, but for graphical intent. As already stated, Figure 1 presents a strong linear relation, with very low variance at almost every clock-gating level. However, the variance increases, i.e. greater box height, with performance decreasing (higher clock-gating levels).
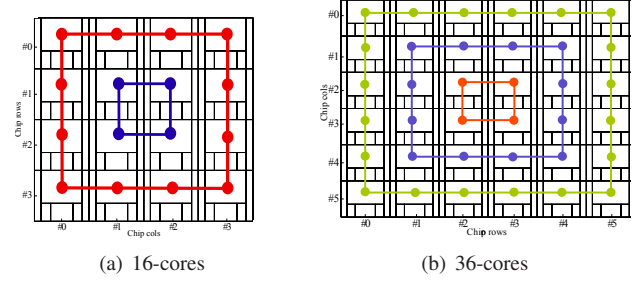
D. Ring-based view in 2D-meshes

The optimization model presented in Section III-A employs the mapping function f to bind tile performance to duty-cycle island it belongs to; however, a suitable analytical formulation of such function has not been provided, yet. This section details the f mapping function formulation to exploit the 2D-mesh topology for thermal/performance trade-off.

The rationale of our mapping function proposal is based on a simple yet effective observation: thermal hotspots are generally located in the centre of a 2D-mesh architecture, independently of the size of the mesh (refer to Section IV for additional details). This fact is tied to the thermal coupling phenomenon: cores surrounded by other cores (as it happens for those located in the centre of the chip) are under the direct influence of core-to-core heat exchange (e.g., through conduction), such that their operating temperature increases up to a point where the thermal management solution is able to dissipate the total system heat. Moreover, the hotspot trend is independent of the mesh size, with the maximum temperature reached by the centre of the chip, and gradually decreasing toward the edges. The only impact of the mesh size is on the maximum operating temperature, with increasing absolute values for aggressive integration made possible by continuous technology scaling. Another key observation relies on the symmetry property of a 2D-mesh thermal map, with respect to all dimensions. Starting from these two observations, the proposed methodology constructs a concentric ring-based set of duty-cycle islands. The concept of *topological ring* is shown in Figure 2 for both 16-cores and 36-cores architectures, with 2 and 3 topological rings respectively. A topological ring is associated with a set of cores in the architecture, and rings are placed concentric each other. Each core belongs to one and only one ring, and cores belonging to the same ring share similar temperature dissipation properties: for instance, all the cores belonging to the outermost topological ring are placed against the chip edge, with direct impact on the way heat is dissipated and temperature is exchanged with package and ambient [24].

IV. EXPERIMENTAL RESULTS

The methodology proposed in Section III is general, while its validity is hereby shown for a reference architecture. In this perspective, we have focused on a real environment scenario

to validate the goodness of the methodology in a real-world context, to demonstrate the practical solution found. Section IV-A details simulation setup and experimental settings, and the steps to assess the proposed methodology. Section IV-B reports and discusses a preliminary analysis on the role of thermal coupling in setting the operating temperature of a multi-core architecture: we will show that the high density of cores in a multi-core architecture makes the central region of the silicon die more spotted to reliability concerns. Section IV-C reports strengthening results obtained on the selected reference architecture. Section IV-D shows how the proposed model can constraint the operating temperature, given a tunable threshold: in reliable designs, this is of utmost relevance in determining the lifetime of the device. Last, temperature/performance trade-off is shown in Section IV-E.

A. Experimental setup and methodology evaluation

We conducted several experiments using a modified version of `GEM5` as an appropriate cycle-accurate simulator (http://gem5.org), a modified version of `McPAT` [17] and `Orion` [13] detailed models for cores and routers power consumption estimates, and the widely used `HotSpot` thermal model [24] to generate chip temperature map. The reference architecture we target is an Alpha21364 network architecture [20], that is used in real Web-servers and Data-centre contexts; commercial examples exist for this kind of architecture, based on the Alpha21264 processor core. We selected and simulated two different architecture configurations, with 16 and 36 cores based on the Alpha21364 architecture. We conducted the experiments with the architecture configuration presented in Table I for typical $45nm$ technology node. Each tile in the network architecture is composed of a single Alpha21264 core, $1.75MB$ local (shared) L2 cache memory and a router to interface to the NoC; its logical architecture is reported in Figure 3 for reference.

We assess the soundness of the proposed methodology within four main steps. First, a set of $500 + 500$ experiments are conducted on both 16-cores and 36-cores architectures to collect representative samples for both temperature and performance related to different duty-cycle levels. Each experiment runs for 2×10^7 instructions per core with a different benchmark mix randomly selected from our representative pool of benchmark suites. We used `WCET` benchmarks from Mälardalen University [7], `SPLASH2` [29] from the University of Delaware, and `MiBench` [8] to cover a broad range of applications, with a mix of integer, floating-point and memory instructions. For each experiment different duty-cycle levels are set for each topological ring in the architecture. Starting from such raw data, we have estimated both the thermal and performance model, described in Section III-B and III-C respectively, using a least squares approach. Then, for a selected set of $10 + 11$ temperature levels, we run the optimization model to obtain duty-cycle levels for each ring to achieve the desired chip temperature.
We run 20 different simulation for each optimized temperature duty-cycle, for a total of 10×20 simulation on 16-cores

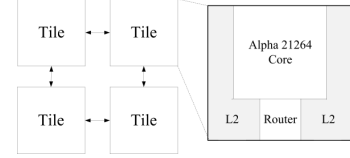| Processor core	3GHz, in-order based on Alpha21264 core
Int-ALU	4 integer ALU functional units
Int-Mult/Div	4 integer multiply/divide functional units
FP-Mult/Div	4 floating-point multiply/divide functional units
L1 cache	64kB 2-way set assoc. split I/D, 2 cycles latency
L2 cache	1.75MB per bank, 8-way associative
Router	2-stage wormhole switched (Garnet network [1])
Topology	2D-mesh based on Alpha21364 network processor
Technology	45nm at 1.1V

Fig. 3. Alpha 21364 tile architecture, adapted from [20].

and 11×20 simulation on 36cores. Last, we compared the maximum simulated chip temperature against the predefined threshold, under the performance level found by the optimization model, as reported in Section IV-D and IV-E respectively.

B. Preliminary analysis on topological rings

The methodology presented in this paper is driven by a ring-based view of the target multi-core chip: the processor floorplan is divided into concentric rings, each ring being composed of a predefined set of tiles. The optimization linear model presented in this paper allocates clock-gating levels to cores, according to their placement (i.e., according to the ring they belong to) and according to the desired optimization (e.g., maximum absolute temperature). The rationale of the ring-based methodology has been sketched in Section III, and it is hereby further detailed with experimental results. Figure 4 shows the temperature profile of the 16-cores processor running different applications. The temperature map has been generated after executing 2×10^7 instructions, and after having collected microarchitecture-level statistics to be passed to the power and thermal models. Two aspects are clear from this scenario: the centre of the die has an higher operating temperature with respect to the edges of the silicon die, even though the power consumption of each single core is comparable. This phenomenon is related to the thermal coupling between adjacent cores, causing the centre of the chip to increase the heat dissipation density, increasing the operating temperature. This phenomenon has been shown to get worse with technology scaling [12], but for two to four-cores architectures only. With more cores integrated in the same silicon die, the problem is exacerbated. From a reliability view-point, higher operating temperatures introduce several problems. The thermal profile from Figure 4 presents some variability while crossing horizontally adjacent cores, and this is due to the L2 caches that are known to be cold spots. Routers, on the other hand, contribute to the higher temperature value between cores that are vertically adjacent in the matrix.
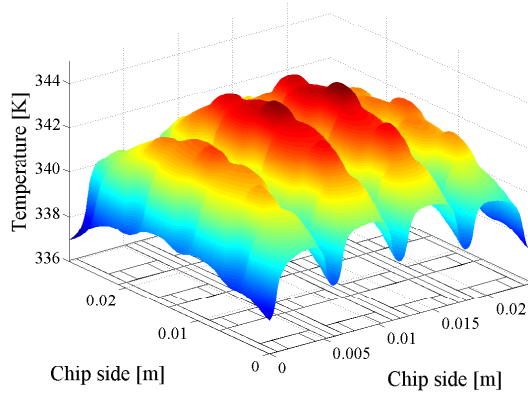
Fig. 4. Temperature profile of a 16-cores processor based on Alpha21364 tiles.

Fig. 5. Temperature profile of a 36-cores processor based on Alpha21364 tiles.

TABLE II
INSTRUCTIONS BREAKDOWN AND POWER CONSUMPTION PROFILE FOR EACH CORE IN THE 16-CORES PROCESSOR, WHOSE TEMPERATURE PROFILE IS GIVEN IN FIGURE 4.

| Core # | Placement | | Instructions | | | Power |
	Row	Col	Int	FP	Mem	[W]
1	0	0	12.8%	6.6%	80.6%	7.132
2	0	1	34.6%	36.5%	28.9%	6.810
3	0	2	64.4%	30.4%	5.3%	7.185
4	0	3	40.8%	29.6%	29.6%	6.717
5	1	0	40.8%	29.6%	29.6%	6.956
6	1	1	64.4%	30.4%	5.3%	6.630
7	1	2	40.8%	29.6%	29.6%	6.849
8	1	3	67.7%	32.3%	0.0%	6.632
9	2	0	66.1%	33.9%	0.0%	6.909
10	2	1	80.7%	19.2%	0.1%	6.790
11	2	2	99.1%	0.9%	0.0%	6.717
12	2	3	68.0%	31.9%	0.0%	6.848
13	3	0	69.4%	30.6%	0.0%	6.626
14	3	1	66.1%	33.9%	0.0%	7.131
15	3	2	34.6%	36.5%	28.9%	7.042
16	3	3	12.8%	6.6%	80.6%	6.790

Fig. 6. Instructions breakdown per core, and total mix for the 36-cores architecture.

C. Preliminary analysis on applications

The validation of the proposed methodology on a real architecture allows to demonstrate the goodness of our solution, giving us the possibility to exploit the architecture itself to strengthen and generalize our results. Experiments show that considering in-order cores organized in a 2D-mesh, allows to provide an optimal solution that is roughly application-independent. In particular, a detailed view of the applications shows different power consumptions, as sketched in Table II; however, such power differences do not greatly impact thermal map, since this is overwhelmed by the thermal coupling effects. Simply put, we can say that for this specific architecture the effect of different workload is negligible compared to the thermal coupling effects. This result, that is extensively supported by experimental data, allows us to cast a single design-time optimization solution in terms of constrained chip temperature and performance level. Such solution is valid for each application mix that is mapped on the multi-core, providing a great design-time optimization result.

The same situation is seen in 36-cores architectures, where the high number of cores pushes temperature toward further
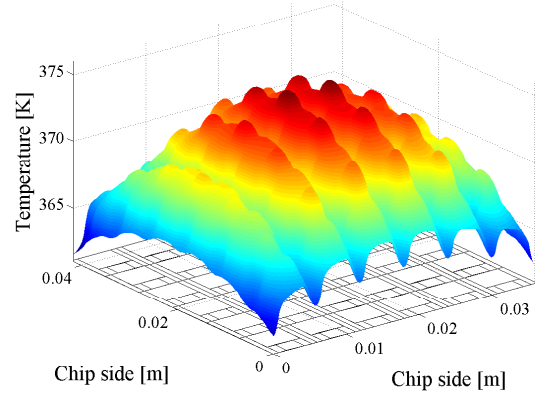
high values. Figure 5 reports the temperature surface of a 36-cores processor, and the relative floorplan. In this case, the workload assignment and instructions breakdown is given in Figure 6, for each core and the total mix.

D. Constraining absolute operating temperature

Reliable designs focus on minimizing operating temperature, to increase the MTTF and reduce the probability of faults. In this work we address hard-faults and not transient ones, and consider two main mechanisms that are known to cause several problems to high-performance processors in scaled technologies [27]: electromigration and stress-migration. We compute the MTTF for these two mechanisms, through the expressions given in Equation 7, taken from [27]: E_{EM} and E_{SM} are the energy activation for electromigration and stress-migration respectively, k is the Boltzmann's constant, T the operating temperature and T_0 the reference temperature for stress-migration (melting temperature), and n is a technology-dependent parameter. We use the values for these parameters as given in [27]. Notice that we consider only the exponential contribution from electromigration, instead of considering the current density since we are assuming to compare results at

TABLE III
MODEL OUTPUT ACCURACY FOR 16-CORES PROCESSOR.

Target threshold temperature [K]	Max. simulated temperature	
	Average [K]	*Variance*
332.0	332.06	0.005
333.0	333.11	0.002
334.0	333.92	0.007
335.0	334.96	0.007
337.0	337.13	0.018
338.0	337.97	0.013
339.0	339.01	0.021
340.0	340.07	0.022
341.0	340.86	0.033
342.0	341.92	0.047

TABLE IV
MODEL OUTPUT ACCURACY FOR 36-CORES PROCESSOR.

Target threshold temperature [K]	Max. simulated temperature	
	Average [K]	*Variance*
350.0	350.13	0.034
352.0	352.48	0.126
354.0	354.32	0.018
356.0	356.15	0.002
358.0	358.20	0.027
360.0	360.16	0.137
362.0	361.95	0.143
364.0	363.93	0.099
366.0	365.82	0.193
368.0	367.74	0.211
370.0	369.67	0.379

different operating temperatures but equal operating conditions (e.g., supply voltage and frequency).

$$MTTF_{EM} \propto exp\left\{\frac{E_{EM}}{k\,T}\right\}$$

$$MTTF_{SM} \propto |T_0 - T|^{-n} \cdot exp\left\{\frac{E_{SM}}{k\,T}\right\} \quad (7)$$

Figure 7 shows the reliability projection of the system for 16-cores and 36-cores processor, while constraining absolute operating temperature. The dotted line shows the theoretical trend of MTTF values with changing temperatures, while circular and diamond markers show the projections ensured by our model: the reliability values are those obtained while employing the optimization model presented in Section III averaged across different runs. The horizontal axis reports the target reliability improvement relative to base case when reliability equals 1. The vertical axis reports the operating temperature required to accommodate such improvement: for example to increase by 40% MTTF caused by electromigration in 16-cores processor, temperature should be diminished to 337K from the 342K base case. Our model ensures that the maximum operating temperature is 337.13K, achieving the expected reliability with an error of less than 1%.

Extensive experimentation has shown a good match between the temperature ensured by the proposed optimization model, and the maximum temperature requirements. Table III and Table IV report the results for the 16-cores and 36-cores processors, respectively. Data is given as an average and variance. Results show a very good match of the computed temperature against the target one, with an average (absolute) error of less than 0.1K and variance in the order of 0.02.

E. Temperature/performance trade-off

The temperature/performance trade-off is depicted in Figure 8 and Figure 9 for 16-cores and 36-cores processor respectively. The plots show a linear relation between the operating temperature and the desired performance. Performance is reported as a percentage over the base-case, when no clock-gating is applied and performance is at 100%. Trade-off linearity is experienced in both architectures, meanwhile presenting a linear relation throughout the entire performance degradation interval from 93% down to 36%. It is worth noticing that
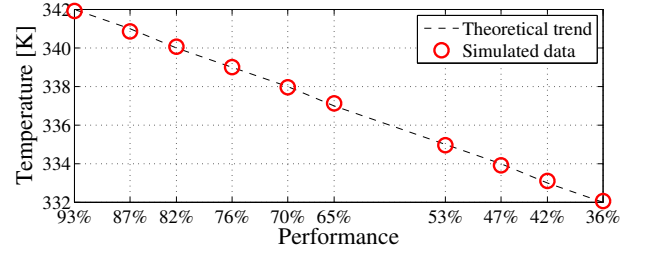
Fig. 8. Temperature/performance trade-off for 16-cores architecture: theoretical trend and simulated data.

Figure 8 and Figure 9 show results for one single generic tile, since the proposed optimization model flattens performance degradation equally on each tile, for each experiment. It is belief of the authors this is a relevant result, since it gives suitable control over operating temperature through a simple relation with respect to core performance.

V. CONCLUSIONS

A joint thermal/performance optimization model has been proposed for design-time optimization of multi-core architectures, as opposed to state-of-the-art Dynamic Thermal Management solutions. Performance and clock-gating have been shown to be linearly related, such that it is possible to use clock-gating as control-knob to seize performance and temperature. Indeed, temperature and performance have been demonstrated to follow a linear relation. The proposed sound and formal model has been used to provide a system-wide optimization framework, focusing on real-world 16-cores and 36-cores processors for high-performance servers and
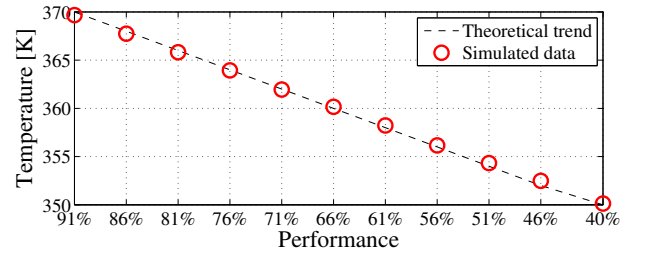
Fig. 9. Temperature/performance trade-off for 36-cores architecture: theoretical trend and simulated data.
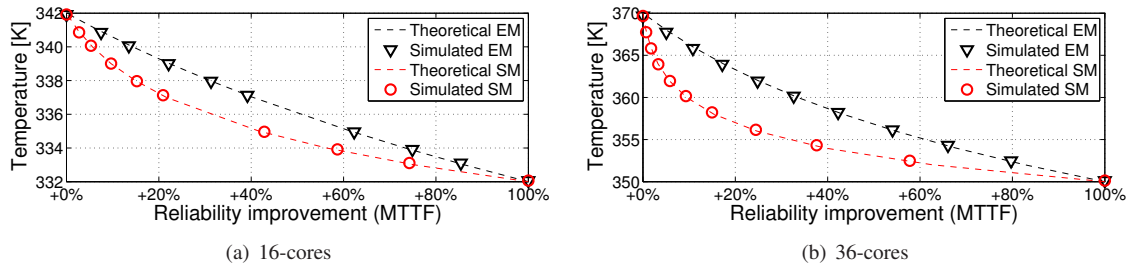
| (a) 16-cores | (b) 36-cores |

Fig. 7. Target reliability improvement against base-case sceanrio (MTTF = 1), and required operating temperatures: theoretical and simulated data.

data-centres. Extensive experimental results have shown the goodness of the proposed optimization model, with respect to 16-cores and 36-cores processors running a predefined set of representative benchmarks. The linear relations have been shown to cover a broad range of temperature and performance situations, such that the proposed methodology is suitable to be employed in real-case scenarios.

REFERENCES

[1] N. Agarwal, T. Krishna, L.-S. Peh, and N. Jha. Garnet: A detailed on-chip network model inside a full-system simulator. In *Performance Analysis of Systems and Software, IEEE International Symposium on*, pages 33 –42, april 2009.

[2] A. M. Azab, P. Ning, and X. Zhang. Sice: a hardware-level strongly isolated computing environment for x86 multi-core platforms. In *ACM Conference on Computer and Communications Security'11*, pages 375–388, 2011.

[3] A. Banerjee, R. Mullins, and S. Moore. A Power and Energy Exploration of Network-on-Chip Architectures. In *NOCS '07*, pages 163–172. IEEE Computer Society, 2007.

[4] A. Bartolini, M. Cacciari, A. Tilli, L. Benini, and M. Gries. A virtual platform environment for exploring power, thermal and reliability management control strategies in high-performance multicores. In *GLSVLSI'10*, pages 311–316, New York, NY, USA, 2010. ACM.

[5] S. Borkar. Thousand core chips: a technology perspective. In *Annual ACM IEEE Design Automation Conference*, 2007.

[6] D. Brooks and M. Martonosi. Dynamic Thermal Management for High-Performance Microprocessors. In *High Performance Computer Architecture. International Symposium on*, 2001.

[7] J. Gustafsson, A. Betts, A. Ermedahl, and B. Lisper. The mälardalen wcet benchmarks - past, present and future. In *Proceedings of the 10th International Workshop on Worst-Case Execution Time Analysis*, July 2010.

[8] M. R. Guthaus, J. S. Ringenberg, D. Ernst, T. M. Austin, T. Mudge, and R. B. Brown. Mibench: A free, commercially representative embedded benchmark suite. In *Proceedings of the Workload Characterization. IEEE International Workshop*, pages 3–14, Washington, DC, USA, 2001.

[9] S. Heo, K. Barr, and K. Asanović. Reducing power density through activity migration. In *Proceedings of the 2003 international symposium on Low power electronics and design*, ISLPED '03, pages 217–222, New York, NY, USA, 2003. ACM.

[10] Y. Hoskote, S. Vangal, A. Singh, N. Borkar, and S. Borkar. A 5-ghz mesh interconnect for a teraflops processor. *Micro, IEEE*, 27(5):51 –61, sept.-oct. 2007.

[11] M.-y. Hsieh, A. Rodrigues, R. Riesen, K. Thompson, and W. Song. A framework for architecture-level power, area, and thermal simulation and its application to network-on-chip design space exploration. *SIGMETRICS Perform. Eval. Rev.*, 38:63–68, March 2011.

[12] M. Janicki, J. H. Collet, A. Louri, and A. Napieralski. Hot spots and core-to-core thermal coupling in future multi-core architectures. In *Annual IEEE SEMI-THERM Symposium*, pages 205–210. IEEE, Feb. 2010.

[13] A. Kahng, B. Li, L.-S. Peh, and K. Samadi. Orion 2.0: A fast and accurate noc power and area model for early-stage design space exploration. In *DATE.*, pages 423 –428, april 2009.

[14] J. S. Kim, M. B. Taylor, J. Miller, and D. Wentzlaff. Energy characterization of a tiled architecture processor with on-chip networks. In *Proceedings of the International Symposium on Low Power Electronics and Design*, pages 424–427, New York, NY, USA, 2003. ACM.

[15] J. Kong, S. Ching, and K. Skadron. Recent thermal management techniques for microprocessors. *ACM Computing Surveys. To appear*, 2012.

[16] C. J. M. Lasance. Thermally driven reliability issues in microelectronic systems: status-quo and challenges. *Microelectronics Reliability*, 43(12):1969–1974, 2003.

[17] S. Li, J. H. Ahn, R. Strong, J. Brockman, D. Tullsen, and N. Jouppi. Mcpat: An integrated power, area, and timing modeling framework for multicore and manycore architectures. In *Microarchitecture. Annual IEEE/ACM International Symposium on*, pages 469 –480, dec. 2009.

[18] A. Majumdar. Helping chips to keep their cool. *Nature Nanotechnology*, (1):214–215, 2009.

[19] M. Monchiero, R. Canal, and A. González. Design space exploration for multicore architectures: a power/performance/thermal view. In *Proceedings of the 20th annual international conference on Supercomputing*, ICS '06, pages 177–186, New York, NY, USA, 2006. ACM.

[20] S. S. Mukherjee, P. Bannon, S. Lang, A. Spink, and D. Webb. The alpha 21364 network architecture. In *Proceedings of the The Ninth Symposium on High Performance Interconnects*, pages 113–, Washington, DC, USA, 2001. IEEE Computer Society.

[21] T. Ormandy. An Empirical Study into the Security Exposure to Hosts of Hostile Virtualized Environments.

[22] M. Pedram and S. Nazarian. Thermal Modeling, Analysis, and Management in VLSI Circuits: Principles and Methods. *Proceedings of the IEEE*, 94(8):1487–1501, August 2006.

[23] M. D. Powell and T. N. Vijaykumar. Resource area dilation to reduce power density in throughput servers. In *Proceedings of the 2007 international symposium on Low power electronics and design*, ISLPED '07, pages 268–273, New York, NY, USA, 2007. ACM.

[24] K. Skadron, M. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan. Temperature-aware microarchitecture. In *Computer Architecture, 2003. Proceedings. 30th Annual International Symposium on*, pages 2 – 13, june 2003.

[25] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan. TSkadronTACO0oemperature-aware microarchitecture: Modeling and implementation. *ACM Transactions on Architecture and Code Optimization (TACO)*, 1(1), 2004.

[26] V. Soteriou, N. Eisley, H. Wang, B. Li, and L.-S. Peh. Polaris: A system-level roadmap for on-chip interconnection networks. In *ICCD 2006.*, pages 134 –141, oct. 2006.

[27] J. Srinivasan, S. Adve, P. Bose, and J. Rivers. The case for lifetime reliability-aware microprocessors. In *Computer Architecture. Annual International Symposium on*, pages 276 – 287, 2004.

[28] J. Srinivasan and S. V. Adve. Predictive dynamic thermal management for multimedia applications. In *ICS '03: Proceedings of the 17th annual international conference on Supercomputing*, 2003.

[29] The modified SPLASH-2 benchmark suite from University of Delaware http://www.capsl.udel.edu/splash/.

[30] I. Yeo, C. C. Liu, and E. J. Kim. Predictive dynamic thermal management for multicore systems. In *45th Annual Design Automation Conference*, 2008.