



Thermal-throttling server: A thermal-aware real-time task scheduling framework for three-dimensional multicore chips

Ting-Hao Tsai, Ya-Shu Chen*

Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan, ROC



ARTICLE INFO

Article history:

Received 20 March 2015

Revised 1 October 2015

Accepted 27 October 2015

Available online 9 November 2015

Keywords:

Thermal-aware

Real-time task scheduling

Three-dimensional multicore chip

ABSTRACT

Three-dimensional (3D) multicore chips have been recently developed to deal with the power consumption and interconnection delay problems of embedded systems; however, thermal management has proven to be challenging due to the heat effect of vertically stacked cores, and the subsequent trade-off that occurs between performance requirements and overheating. In this paper we propose a novel thermal-aware real-time scheduling framework for 3D multicore chips to achieve an effective trade-off between system temperature and task schedulability for dynamic workloads. A thermal-throttling server is first proposed to adjust the heat generated by task executions, and a thermal-throttling dispatcher is then presented to enable thermal-awareness in well-known real-time dispatchers. An admission control is subsequently derived to ensure that all task executions satisfy the thermal and timing constraints. Lastly, a series of extensive simulations are carried out, with encouraging results in terms of schedulability and the prevention of overheating.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The Multicore System-on-a-Chip design has been widely used in embedded systems due to the high computational requirements of many modern applications. As the number of cores and components in a system increases, so do the power consumption and global interconnect delay to a significant degree. Three-dimensional (3D) stack-based homogeneous multicore designs (Healy et al., 2010; Dreslinski et al., 2013) have been developed to increase transistor density and to improve performance through vertically stacking multiple silicon layers connected by through silicon vias (TSVs) (Cong et al., 2011). The stacked structure, however, causes a higher power density and poorer heat dissipation, which exacerbates the problem of thermal hot spots. Higher on-chip temperatures impact reliability and degrade performance. As such, the trade-off between system performance and temperature has become a critical problem for 3D multicore chips.

Numerous studies have explored thermal-aware scheduling in 3D multicore chips (Coskun et al., 2009; Hameed et al., 2011; Kang et al., 2011; Li et al., 2013; Liu et al., 2010; Lung et al., 2011; Tsai and Chen, 2012; Zhou et al., 2010a; Liao et al., 2015). Ignoring timing constraints, some studies (Coskun et al., 2009; Liu et al., 2010; Zhou et al., 2010a) have utilized thermal-aware task allocations to balance the peak temperatures of each core in order to minimize the peak

temperature; others (Hameed et al., 2011; Lung et al., 2011) have proposed algorithms to minimize the number of hot spots or maximize the throughput. Some studies (Kang et al., 2011; Liao et al., 2015) have applied dynamic voltage scaling to satisfy the thermal constraint. Li et al. (2013) considered the timing constraint for the dependence tasks and proposed a rotation scheduling to minimize the peak temperature. Tsai and Chen (2012) proposed an algorithm to maximize the system utilization under a thermal constraint. To reduce the product size and total cost of the embedded system, thermal constraints would be satisfied using thermal-aware task scheduling instead of additional cooling devices. Liao et al. (2015) assigned tasks according to the nondecreasing order of the peak power and applied the dynamic voltage frequency scaling (DVFS) technique to cores by considering the thermal constraint. For 2D multicore chips, some studies (Chen et al., 2009a; Fisher et al., 2009; Huang et al., 2014) proposed thermal-aware scheduling under thermal/timing constraint(s), and Quan and Chaturvedi (2010) proposed a scheduling utilization bound that took thermal factors into account. Different from the thermal behavior in 2D multicore chips, the thermal behavior in 3D multicore chips needs to take account of the thermal conduction in the vertically aligned cores. The thermal conduction between horizontally adjacent cores is equalization (i.e., the temperature among cores is balanced because heat is dissipated from a high temperature core to a low temperature core); however, the thermal conduction between vertically aligned cores leads to heat accumulation (i.e., the temperature of the core in the upper layer cannot be lower than that of the vertically aligned core in the lower layer). It results in more complicated task dispatching for trade-off the system temperature and

* Corresponding author. Tel.: +886 27376702.

E-mail address: d10007401@mail.ntust.edu.tw, yschen@mail.ntust.edu.tw (Y.-S. Chen).

utilization. Very few thermal-aware real-time scheduling studies of 3D multicore chips have focused on system schedulability in light of the peak temperature constraint.

To investigate the schedulability in the presence of the peak temperature constraint in 3D multicore chips, a partition-based thermal-throttling server was proposed in Tsai and Chen (2014). In contrast to the hardware throttling mechanism, a thermal-throttling server manages the thermal budget of each core by reserving appropriate cooling utilization without violating the timing constraint. In this study, we extended the thermal-throttling server to a global dispatcher and also devised a method to determine the thermal size of each core for meeting the run-time thermal constraint. Admission control is also presented for a given task set to improve the system robustness. The thermal management capabilities of global and partition schemas were evaluated in the absence admission control, and the derived schedulability of admission control under global and partition schemas and different thermal size assignments were compared to provide system designers with further insight.

This study is motivated by a desire to investigate the complexities raised by the trade-off that exists between system performance and temperature in 3D multicore chips. Our contributions are as follows:

- A novel thermal-aware multicore real-time scheduling framework is proposed to dynamically manage the run-time system temperature and provide a quality of service guarantee for 3D stacked multicore chips.
- To deal with the thermal conduction from vertically aligned cores in 3D multicore chips, thermal size assignment is proposed to manage the heat generated by task executions. The thermal management effect on system performance is also evaluated.
- In contrast to off-line optimization approaches, a thermal-throttling server is proposed to determine and change the core's execution frequency for run-time temperature management based on the given thermal size and incoming task information.
- A thermal-throttling dispatcher is presented as an extension of well-known real-time dispatchers, i.e., First-Fit and Global, for run-time thermal-aware task dispatching and for dynamically switching the thermal size assignment policy with dynamic workload.
- The thermal-aware admission control is proposed to ensure that all tasks meet the timing and thermal constraints, from the perspective of single task, single core, and a task set with multicore, enabling thermal awareness in all real-time schedulers.

The rest of this paper is organized as follows: Section 2 presents the system model; Section 3 proposes a thermal-aware real-time task scheduling framework for a 3D multicore system, including thermal size assignment, thermal-throttling server, thermal-throttling dispatcher, and admission control; Section 4 details the performance evaluation; Section 5 presents our conclusions.

2. System model

This study considers a 3D homogenous multicore chip, as shown in Fig. 1, and the chip consists of a set of active silicon layers $\mathbf{L} = \{L_1, L_2, \dots, L_h\}$. Each active silicon layer has one or more cores, and a thermal interface material (TIM) is used between each silicon layer to provide efficient heat transfer. The vertically aligned cores are modeled as a *Stack*, and they communicate each other by TSVs. In order to remove heat, one side of the chip is soldered to the heat sink. The bottom layer closest to the heat sink is L_1 , and the cores on L_1 are *bottom cores*. The top layer is the layer farthest from the heat sink, and the cores located in this layer are *top cores* (Zhou et al., 2010a).

The proposed system consists of a set of independent periodic tasks $\mathbf{\Gamma} = \{\tau_1, \tau_2, \dots, \tau_m\}$, where m is the number of tasks in $\mathbf{\Gamma}$. Each task τ_i in the system is associated with a worst-case execution core

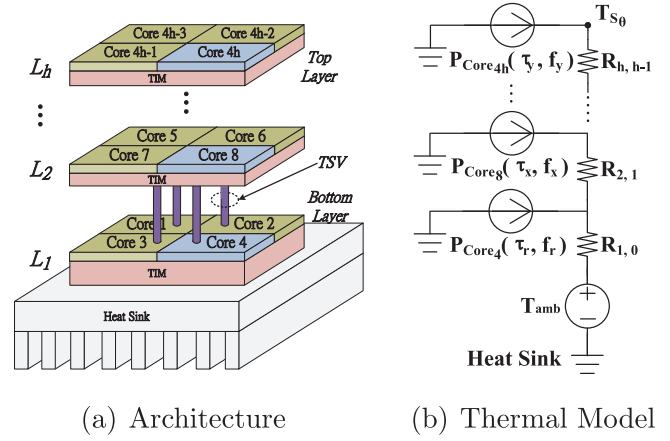


Fig. 1. 3D multicore chip.

cycle c_i (Wilhelm et al., 2008), a power consumption $P_{Core_j}(\tau_i, f_{max})$, and a relative deadline (period) d_i . The utilization u_i of a task τ_i is the ratio of the worst-case execution time to the relative deadline, when the task is executed at the maximum frequency of the system (i.e., $\frac{c_i}{f_{max} \times d_i}$) (Liu and Layland, 1973). The number of core cycles executed in a time interval depends on the core frequency. In other words, the computation time for a task τ_i execution at the frequency f_i is $\frac{c_i}{f_i}$.

Each core $Core_j$ in the system has the same computing resources and an independent dynamic voltage/frequency scaling (DVFS) capabilities. The power consumption function of each core running a task τ_i can be divided into dynamic power and leakage power consumptions, $P_{Core_j}(\tau_i, f_i) = P_{dyn,j}(\tau_i, f_i) + P_{leak,j}$. The dynamic power consumption $P_{dyn,j}(\tau_i, f_i)$ can be modeled as a convex function of the core frequency (Rabaey et al., 2002): $P_{Core_j}(\tau_i, f_i) = \frac{1}{2} C_{ef,\tau_i} V_{dd}^2 f_i$, where C_{ef,τ_i} , V_{dd} , and κ denote the effective switch capacitance when executing the task τ_i , the supply voltage, and the hardware-design-specific constant, respectively. The above equation shows that $P \propto f_i^\rho$ (Zhu, 2006; Xu et al., 2005); we assume that ρ is 2 in this paper,¹ and assumes that the core can operate at any frequency f_i in a given range $[f_{min}, f_{max}]$, where f_{min} is larger than the lowest frequency required to eliminate the effect of leakage power consumption (Jejurikar et al., 2004). Further extensions of this work for processors with discrete available frequencies can be done using approaches similar to those reported in Huang et al. (2014), Chaturvedi et al. (2010). The leakage power consumption $P_{leak,j}$ can be modeled as $P_{leak,j} = h + \varsigma T_j$, where T_j is the absolute temperature on $Core_j$ and h and ς both are constants (Chen et al., 2009b; Liao et al., 2005; Chantem et al., 2011; Liu et al., 2007).

In a single core, according to Fourier heat flow analysis and the RC model (Skadron et al., 2002), the approximate temperature of each $Core_j$ at time t can be estimated as:

$$T_{Core_j}(t) = P_{Core_j}(\tau_i, f_i) \times R_j + T_{amb} + (T_{init} - (P_{Core_j}(\tau_i, f_i) \times R_j + T_{amb}))e^{-t/R_j C_j} \quad (1)$$

where T_{init} is the initial temperature of the core, $P_{Core_j}(\tau_i, f_i)$, R_j , C_j are the power consumption, the thermal resistance, and the thermal capacitance of $Core_j$. According to Eq. (1), the heat of $Core_j$ caused by a task τ_i at a frequency f_i is no larger than $P_{Core_j}(\tau_i, f_i) \times R_j$. The accuracy of this function depends on the accuracy of the power consumption estimate. In this study, the peak power was considered to avoid system overheating.

In a 3D multicore chip, thermal coupling of the vertically adjacent cores leads to a larger thermal influence among them compared with

¹ The proposed approach is also applied for varying the value of ρ (e.g., $\rho \geq 2$).

Download English Version:

<https://daneshyari.com/en/article/458335>

Download Persian Version:

<https://daneshyari.com/article/458335>

[Daneshyari.com](https://daneshyari.com)