An Online Thermal-Constrained Task Scheduler for 3D Multi-Core Processors

Chien-Hui Liao
Inst. of Communications Engineering
National Chiao Tung University
Hsinchu, Taiwan 30010
Email: liangel.cm97g@g2.nctu.edu.tw

Charles H.-P. Wen
Electrical and Computer Engineering
National Chiao Tung University
Hsinchu, Taiwan 30010
Email: opwen@g2.nctu.edu.tw

Krishnendu Chakrabarty
Electrical and Computer Engineering
Duke University
Durham, NC 27708
Email: krish@ee.duke.edu

Abstract—**Hotspots occur frequently in 3D multi-core processors (3D-MCPs) and they can adversely impact system reliability and lifetime. Moreover, frequent occurrences of hotspots lead to more dynamic voltage and frequency scaling (DVFS), resulting in degraded throughput. Therefore, a new thermal-constrained task scheduler based on thermal-pattern-aware voltage assignment (TPAVA) is proposed in this paper. By analyzing temperature profiles of different voltage assignments, TPAVA pre-emptively assigns different operating-voltage levels to cores for reducing temperature increase in 3D-MCPs. Moreover, the proposed task scheduler integrates a vertical-grouping voltage scaling (VGVS) strategy that considers thermal correlation in 3D-MCPs. Experimental results show that, compared with two previous methods, the proposed task scheduler can respectively lower hotspot occurrences by 47.13% and 53.91%, and improve throughput by 6.50% and 32.06%. As a result, TPAVA and VGVS are effectively for reducing occurrences of hotspots and optimizing throughput for 3D-MCPs under thermal constraints.**

I. Introduction

3D integration offers several advantages, such as shorter interconnect delay and better performance [1], over traditional two-dimensional (2D) chip implementation. However, 3D multi-core processors (3D-MCPs) are likely to exhibit severe thermal problems due to high power density and lack of heat dissipation paths [2]. For example, the peak temperature of a 3D two-layer eight-core processor may increase by $20\,^{\circ}\mathrm{C}$, when compared with that of a 2D eight-core processor [3]. Moreover, the temperature gradient of a 3D two-layer eight-core processor is also much steeper than that of a 2D eight-core processor [4]. Due to the dramatic temperature increase and non-uniform temperature distribution in 3D-MCPs, hotspots occur more often and cause problems such as severe thermal stress, worse system reliability, shorter lifetime, slower switching speed, and higher cooling cost.

To handle such serious thermal problems in 3D-MCPs, many previous methods such as dynamic thermal management (DTM), microchannel liquid cooling [5], floorplanning [3] [6] and thermal TSV [7] have been proposed at the architectural and the physical-design levels. In particular, a specific DTM techniques, namely dynamic voltage and frequency scaling (DVFS), has received much attention recently to deal with the thermal crisis [8]. DVFS can control the temperature of 3D-MCPs efficiently by lowering operating voltages of cores, but it is unfortunately accompanied by degraded throughput. Therefore, thermal constraints and the performance (in terms of the throughput) should be considered simultaneously in any effective DVFS solution.

When frequent hotspots occur, more DVFS actions are triggered, causing more degradation in the throughput. Therefore, many OS-level task-scheduling algorithms [8][9][10][11][12][13][14] focused on thermal-aware task allocation have been proposed to reduce hotspots, and thereby maximize the throughput under thermal constraints. In addition, most of today's processors are integrated with thermal sensors [15] for tracking the operating temperature of cores. Hence, the task scheduler can readily detect processor overheating through these thermal sensors.

In most previous task schedulers as shown in Figure 1(a), all cores operate at the highest voltage level at the beginning, and then DVFS is triggered on the overheated core when the temperature approaches a preset threshold. However, two promising directions have not yet been explored in prior work. First, different strategies of voltage scaling may result in different effects on cooling in 3D-MCPs. Second, initial voltage assignment may have a great impact on temperature increase in 3D-MCPs. Therefore, we develop a new voltage-scaling strategy to deal with thermal emergency, and propose a voltage-assignment strategy for reducing temperature increase in 3D-MCPs.

In this paper, a new thermal-constrained task scheduler based on a thermal-pattern-aware voltage assignment (TPAVA) for 3D-MCPs is proposed. It consists of an on-line allocation strategy and a new voltage-scaling strategy to maximize throughput and minimize occurrences of hotspots. First, the thermal behavior of 3D-MCPs under different voltage levels is analyzed to generate temperature profiles of all cores under different voltage levels. According to these temperature profiles, TPAVA assigns different initial operating-voltage levels to all cores for reducing the temperature increase in 3D-MCPs. In addition, two 3D-MCP thermal behaviors [8] on vertically-

The work of C.-H. Liao was supported by the Ministry of Science and Technology, Taiwan, through a research scholarship with Duke University, Durham, NC, USA, from 2013 to 2014, under Grant 102-2917-I-009-035.

978-3-9815370-4-8/DATE15/©2015 EDAA

aligned cores and core-to-heat-sink distance are utilized by our on-line allocation and vertical-grouping voltage scaling (VGVS) strategies to reduce hotspots in 3D-MCPs.

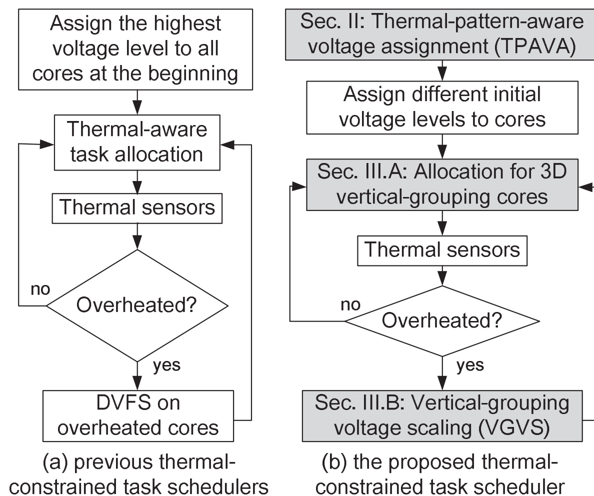

Fig. 1. Comparison between previous thermal-constrained task schedulers and the proposed thermal-constrained task scheduler

Since frequent occurrences of hotspots not only influences throughput, but also reduces system lifetime, the reduction of hotspots becomes a critical target of a 3D-MCP task scheduler. Experimental results show that our proposed thermal-constrained task scheduler outperforms two previous methods, STSC [8] and Adapt3D [11], by 47.13% and 53.91% in terms of hotspot reduction. In addition, our proposed task scheduler can achieve higher throughput by 6.50% and 32.06% on average than the STSC and Adapt3D task schedulers, respectively. As a result, the proposed task scheduler significantly improves upon previous work by simultaneously reducing hotspots and optimizing throughput for 3D-MCPs under thermal constraints.

Figure 1(b) shows the key components of this work. Furthermore, the main contributions of this work are summarized below:

- Temperature profiles of 3D-MCPs under different voltage levels are first analyzed in this paper. Based on these temperature profiles, a novel thermal-pattern-aware voltage assignment (TPAVA) strategy is proposed to pre-emptively assign different voltage levels to all cores for reducing the temperature increase in 3D-MCPs.
- By considering the strong thermal correlation among vertically-aligned cores, a new vertical-grouping voltage scaling (VGVS) strategy is proposed for cooling 3D-MCPs effectively. The VGVS strategy triggers DVFS on not only the overheated core, but also all vertically-aligned cores of the overheated core.

The rest of this paper is organized as follows. In Section II, temperature profiles of 3D-MCPs under different voltage levels are analyzed. According to these temperature profiles, a thermal-pattern-aware voltage assignment (TPAVA) strategy is proposed. Next, a new thermal-constrained task scheduler with TPAVA is elaborated in Section III. Two core techniques including on-line allocation for vertical-grouping cores and vertical-grouping voltage scaling (VGVS) are introduced. In Section IV, experimental results highlight the better performance of the proposed task scheduler with rest to throughput improvement and hotspot reduction under thermal constraints, compared to two previous methods [8][11]. Finally, Section IV concludes this paper.

II. Thermal-Pattern-Aware Voltage Assignment (TPAVA)

Previous OS-level task-scheduling algorithms focused on thermal-aware task allocation have been proposed to optimize the throughput under thermal constraints. Coskun et. al. [9] proposed a temperature-aware allocation strategy, Coolest, which assigned the task to the coolest core one by one according to the current temperature. In [10], instead of migrating an overheated task to the coolest core, the scheduler migrated the task to the core that takes the longest period to reach the preset temperature threshold. Adapt3D [11] used the temperature history of cores to compute the probabilities of assigning jobs to cores. An approach in [12], proposed a fast task-allocation algorithm through rewriting thermal equations. Zhou et. al. [8] proposed a Super-Task-to-Super-Core strategy, which defined super tasks and allocated super tasks to super cores.

In most previous work, all cores operate at the highest voltage level at the beginning, and then DVFS is triggered for the overheated core when the temperature approaches a preset threshold. However, voltage assignment is found to have a great impact on temperature increase in 3D-MCPs. Therefore, we start from analyzing thermal behavior of 3D-MCPs under different voltage levels to generate temperature profiles of all cores under different voltage levels. Then, based on the temperature profiles, a Thermal-Pattern-Aware Voltage Assignment (TPAVA) is proposed to assign different initial operating-voltage levels to all cores for reducing temperature increase in 3D-MCPs.

A typical 3D-MCP schema from HotSpot5.0 [16][17] is used in this work, which includes printed circuit board (PCB), heat sink, heat spreader, thermal interface material (TIM), dies, and other components. Note that the heat sink is placed on top of the 3D-MCP in this work. Thermal simulation using HotSpot5.0 is typically used to compute the temperature for checking whether thermal constraint is satisfied. HotSpot5.0 has an automated thermal model for 3D-MCPs and is built upon the well-known electrical-and-thermal duality [18].

The thermal behaviors of 3D MCPs are quite different from those of 2D MCPs [2]. For a thermal-aware task-scheduling algorithm, adequate understanding of 3D-MCP thermal behaviors is important to ensure effectiveness. Two important 3D-MCP thermal behaviors are found in [8][16]: (1) vertically-aligned cores are strongly thermal-correlated; (2) cores farther from the heat sink are hotter. They are utilized during the design of task scheduler for satisfying thermal constraints and optimizing throughput.

In addition, the temperature-increase rate (dT/dt) of a core can be computed using Equation (1) where T_s is the

steady-state temperature, $T(t)$ is the current temperature and b is a thermal parameter of 3D-MCPs [19]. By solving Equation (1), the current temperature can be obtained using Equation (2) where $T(0)$ is the initial temperature of 3D-MCPs. The temperature-increase rate is proportional to the difference between the steady-state temperature and the current temperature. Therefore, if the steady-state temperature is higher than the current temperature, the temperature will increase exponentially. On the other hand, if the steady-state temperature is lower than the current temperature, the temperature will decrease exponentially.

$$\frac{dT}{dt} = b \times (T_s - T(t)) \quad (1)$$

$$T(t) = T_s - (T_s - T(0)) \times e^{-bt} \quad (2)$$

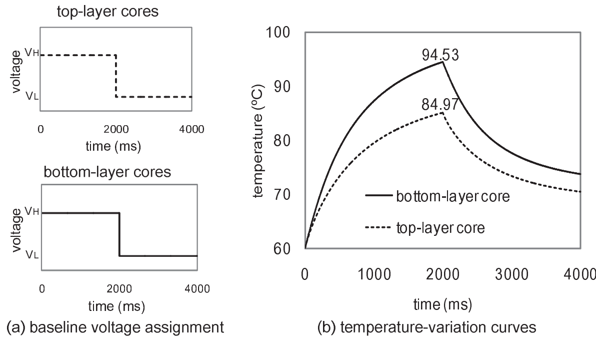

Fig. 2. An example of temperature-variation curves under the baseline voltage assignment

An example in Figure 2 illustrates the temperature-increase curves and temperature-decrease curves of a 3D two-layer processor under different voltage levels. In the example, all cores are assigned the same tasks. We show the influences of different voltage assignments on temperature variations. In Figure 2(a), all cores are first assigned the higher voltage V_H from 0 ms to 2000 ms, and then assigned the lower voltage V_L from 2000 ms to 4000 ms. Figure 2(b) shows the thermal-simulation result from HotSpot5.0 [17][16]. In the example, the initial temperature of the 3D-MCP is $60\,^{\circ}\mathrm{C}$, and the steady-state temperatures with V_H and V_L are $133.97\,^{\circ}\mathrm{C}$ and $69.91\,^{\circ}\mathrm{C}$, respectively. Therefore, the temperature first increases to $94.53\,^{\circ}\mathrm{C}$, and then at 2000 ms, the temperature turns to decrease because the steady-state temperature with V_L is lower than $94.53\,^{\circ}\mathrm{C}$. In this case, the peak temperature is $94.53\,^{\circ}\mathrm{C}$.

Furthermore, to illustrate the influence of voltage assignments on temperature increase, we present examples with different voltage-assignment strategies, namely vertically-reversed voltage assignment and extreme voltage assignment. Considering strong thermal correlation among vertically-aligned cores, the vertically-reversed voltage-assignment strategy reverses the voltage of top-layer cores with the voltage of bottom-layer cores. In Figure 3(a), the vertically-reversed voltage-assignment strategy is applied to the same case in Figure 2. From 0 ms to 2000 ms, the bottom-layer cores are

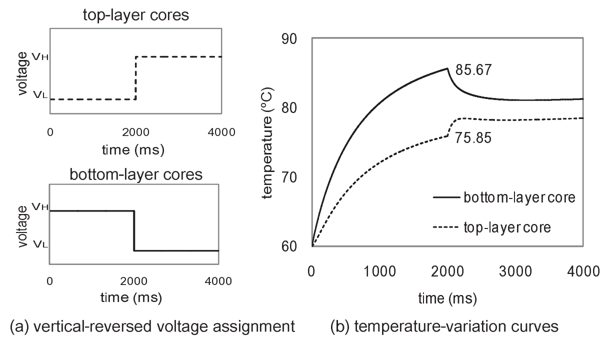

Fig. 3. An example of temperature-variation curves obtained using the vertically-reversed voltage-assignment strategy

assigned V_H, but the top-layer cores are assigned V_L. Then, from 2000 ms to 4000 ms, the bottom-layer cores are assigned V_L, but the top-layer cores are assigned V_H. As a result, in Figure 3(b), the peak temperature is reduced to $85.67\,^{\circ}\mathrm{C}$ by the vertically-reversed voltage-assignment strategy.

Moreover, considering that cores farther from the heat sink are hotter, the extreme voltage-assignment strategy assigns different voltages to cores on different layers. In Figure 4, extreme voltage-assignment strategy assigns V_H to top-layer cores and assigns V_L to bottom-layer cores from 0 ms to 4000 ms. As a result, the peak temperature resulting from the extreme voltage-assignment strategy is only $80.99\,^{\circ}\mathrm{C}$, which is much lower than the peak temperature in the previous example. Note that the throughputs of the three examples in Figure 2-4 are equal because the periods of triggering DVFS (using the lower voltage V_L) are all the same. Through these examples, we have demonstrated the influences of voltage-assignment strategies on temperature increase.

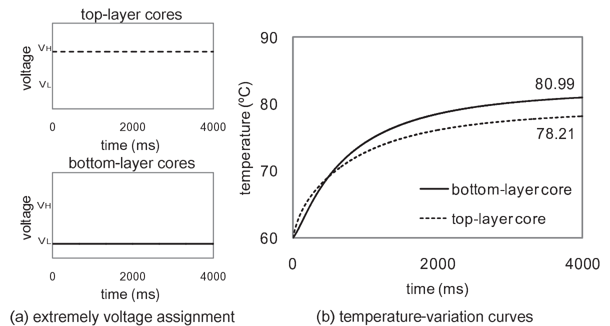

Fig. 4. An example of temperature-variation curves obtained using the extreme voltage-assignment strategy

In addition, Figure 5 compares the peak temperatures of three voltage-assignment strategies on 3D four-layer MCPs. The extreme voltage-assignment strategy assigns V_H to higher-layer ($layer3$ and $layer4$) cores and assigns V_L to lower-layer ($layer1$ and $layer2$) cores. Compared to the baseline voltage assignment, the peak temperature can be reduced by $20\,^{\circ}\mathrm{C}$ with the extreme voltage-assignment strategy. Again, the effectiveness of reducing the temperature increase by extreme voltage-assignment strategy is demonstrated. Therefore, based on the temperature-profile analysis, the thermal-pattern-aware

voltage assignment (TPAVA) is proposed. TPAVA assigns the lower voltage to lower-layer cores and assigns the higher voltage to higher-layer cores for reducing temperature increase effectively. Note that the task scheduler integrated with TPAVA will keep the voltage level of cores equal to or lower than the voltage level assigned by TPAVA for all time. The effectiveness of TPAVA for reducing hotspots and improving throughput will be shown in Section IV.

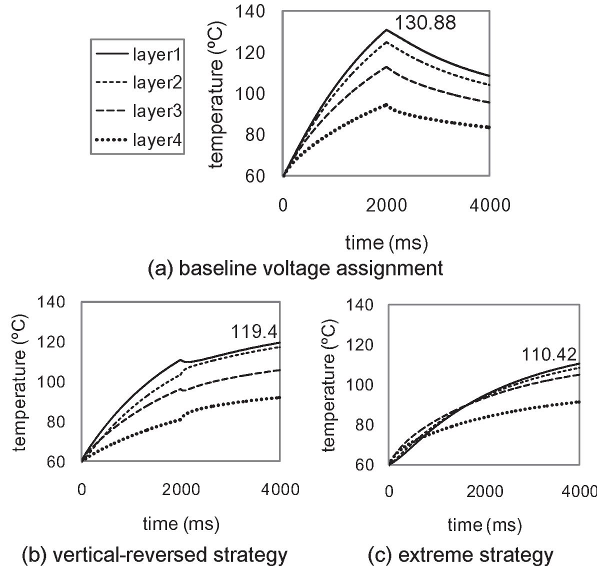

Fig. 5. Comparison of the temperature increase for different voltage-assignment strategies in 3D four-layer MCPs

III. On-Line Allocation and Voltage Scaling

A new thermal-constrained task scheduler based on TPAVA is proposed to optimize throughput and reduce occurrences of hotspots. The proposed thermal-constrained task scheduler includes two techniques: on-line allocation of vertical-grouping cores and vertical-grouping voltage scaling (VGVS). Since more hotspots occur, more DVFS actions are triggered and this leads to degraded throughput. Therefore, both on-line allocation and VGVS consider thermal behaviors of 3D-MCPs to reduce hotspots as much as possible and decrease the temperature of overheated 3D-MCPs more effectively.

A. On-Line Allocation for Vertical-Grouping Cores

Considering strong thermal correlation among vertically-aligned cores, Zhou et. al. [8] proposed a Super-Task-to-Super-Core (STSC) strategy, which defined super tasks and super cores. Since vertically-aligned cores have strong thermal correlations that should be considered together, STSC groups vertically-aligned cores as a super core. For example, a 3D two-layer eight-core processor can be divided to four super cores, and each super core consists of two vertically-aligned cores. Furthermore, STSC allocated a super task, i.e., a set of tasks, to a super core. For a 3D-MCP with N layers, a super task consists of N tasks. Therefore, the three-dimensional allocation problem can be reduced to two-dimensional allocation problem by the STSC strategy.

Figure 6 shows an example of the STSC allocation strategy for a 3D two-layer eight-core processor. At the beginning of each scheduling interval, power consumption of tasks is sorted in an increasing order as shown in Figure 6(a). For balancing the power distribution among super cores, the total power consumption of super tasks should be balanced. Therefore, the largest-power task and the smallest-power task will be merged into one super task. Furthermore, since cores farther from the heat sink are hotter, the larger-power tasks will be assigned to the top-layer cores. In Figure 6(b), the super tasks are sorted in increasing order of total power consumption. Then, super cores are sorted in decreasing order of the temperature summations as shown in Figure 6(c). Finally, in Figure 6(d), the super tasks are allocated to the super cores. In addition, Figure 6(e) shows that the peak temperature of the 3D-MCP becomes $108\,^{\circ}\mathrm{C}$ after five seconds when the STSC allocation strategy is used.

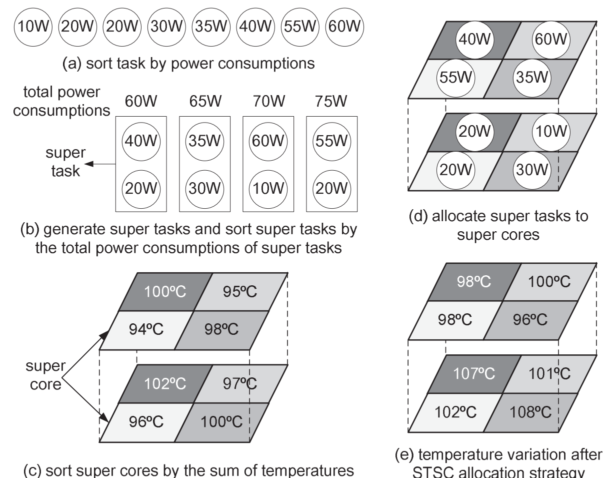

Fig. 6. An example of the STSC allocation strategy for a 3D two-layer eight-cores processor [8]

In Figure 6, because the super tasks are sorted by the total power consumption, the task with 30 W power consumption is allocated to a core with $100\,^{\circ}\mathrm{C}$. However, such allocation makes the core running with 30 W task heated to $108\,^{\circ}\mathrm{C}$ after five seconds. Therefore, considering that the temperature of bottom-layer cores dominates the peak temperature of the entire system, unlike the STSC strategy, our on-line allocation modifies the super-task sorting strategy. In our on-line allocation strategy, super tasks are sorted in increasing order of the bottom-layer-task power consumption. Figure 7 shows the same example for the proposed on-line allocation strategy. Figure 7(a) shows the different order of super tasks in the increasing order of bottom-layer-task power consumption. Figure 7(b) shows our new allocation and Figure 7(c) shows that the peak temperature of the 3D-MCP becomes $105\,^{\circ}\mathrm{C}$ after five seconds. As a result, our on-line allocation strategy outperforms the STSC strategy on peak-temperature reduction by $3\,^{\circ}\mathrm{C}$.

B. Vertical-Grouping Voltage Scaling (VGVS)

The other key component of our task scheduler is how to decrease 3D-MCP temperature whenever the temperature ex-

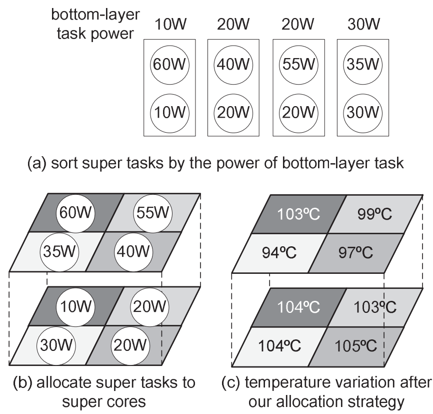

Fig. 7. An example of the proposed allocation strategy for a 3D two-layer eight-cores processor

ceeds the preset threshold. Conventionally, traditional voltage-scaling strategy triggers DVFS only on the overheated core as shown in Figure 8(a). However, for cooling 3D-MCPs faster, two voltage-scaling strategies, club-grouping voltage scaling and vertical-grouping voltage scaling (VGVS), are proposed. Figure 8(b) illustrates the club-grouping voltage-scaling strategy, which triggers DVFS not only on the overheated core, but also on all cores adjacent to the overheated core. In Figure 8(c), the VGVS strategy triggers DVFS on the overheated core as well as vertically-aligned cores of the overheated core.

Given a 3D-MCP that is overheated to $95\,^{\circ}\mathrm{C}$, the comparison of the temperature decrease for three voltage-scaling strategies is shown in Figure 8(d). The club-grouping voltage-scaling strategy and the VGVS strategy result in faster cooling than the traditional voltage-scaling strategy. In addition, the temperature-decrease curves of the club-grouping voltage-scaling strategy and the VGVS strategy are overlapped. However, since the club-grouping voltage-scaling strategy triggers more DVFS on more cores, leading to more degraded throughput, the VGVS strategy is a better choice than the club-grouping voltage-scaling strategy. Furthermore, compared to the traditional voltage-scaling strategy, although the VGVS strategy may trigger more DVFS on all vertically-aligned cores, it cools 3D-MCPs faster and thus achieves better throughput.

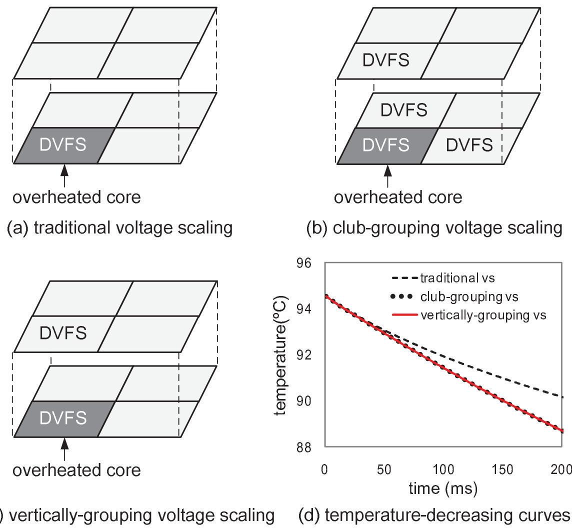

Fig. 8. Comparison of the temperature decrease for three different voltage-scaling strategies

IV. Experimental Result

Our experiments are conducted on SPEC CPU 2006 benchmarks on different architectures of 3D MCPs including 8-core two-layer, 12-core three-layer and 16-core four-layer 3D MCPs, respectively. Each benchmark was ran for one hour to trace utilization of each hardware thread by $mpstat$ (a type of Linux commend) at every second. Then, the utilization-percentage profile was used to transform the SPEC CPU2006 benchmarks into task graphs with power-consumption information [11]. The task-scheduling interval is 100 ms and the thread-migration overhead is 1 ms [11]. In addition, the current temperature is checked every 10 ms ($1/10$ scheduling interval) for ensuring the thermal constraint is satisfied. Grid-type thermal simulation using HotSpot5.0 [17] [16] with a multi-core version of the Alpha 21264 floorplan (each core area is 5 mm $\times$ 5 mm) is performed to compute temperatures of cores after task scheduling. The default parameters of HotSpot5.0 for 3D MCPs are used in this work.

In this paper, we assume that cores can be run at different voltages and frequencies. The number of the operational voltage levels is determined in practice by the designers and the manufacturing technology. The highest voltage is 1.3 V and the highest frequency is 3 GHz [11]. For simplicity, only three voltage-and-frequency levels (i.e. 1.30 V-3.0 GHz, 1.17 V-2.7 GHz and 1.05 V-2.4 GHz) are used here. The power model from [11] is used to perform voltage scaling. The dynamic-power consumption at the highest voltage level of an Alpha 21264 processor ranges between 40 W and 70 W [8][20] and the dynamic-power consumption at the lower voltage level is proportional to $f \cdot v^2$. In addition, a temperature-dependent leakage power model within HotSpot5.0 [17] [16] is considered. Moreover, the delay of tasks running at the lower voltage level is proportional to $1/f$. The DVFS overhead is assumed to 30 μs [8].

In general, considering the reliability of systems, the peak temperature of cores should be kept below $85\,^{\circ}\mathrm{C}{-}110\,^{\circ}\mathrm{C}$ [21]. Therefore, in the experiments, the thermal constraint sets the peak temperature below $85\,^{\circ}\mathrm{C}$. Moreover, in our experiments, the effectiveness of hotspot reduction by our task scheduler is evaluated, where a hotspot is defined to occur when the temperature exceeds $85\,^{\circ}\mathrm{C}$ in the 3D-MCP.

Table I shows the comparison in terms of hotspot reduction rate (HRR) and throughput improvement rate (TIR) between the proposed task scheduler and the STSC task scheduler [8] under thermal constraints. HRR and TIR are computed using Equation (3) and Equation (4), respectively. The first column shows the number of layers on 3D-MCPs. The second to the seventh columns show the HRR and the TIR by our on-line allocation strategy with TPAVA, our on-line allocation strategy

with VGVS, and our on-line allocation strategy with both VGVS and TPAVA, respectively. As a result, compared to the STSC task scheduler, our on-line allocation strategy with VGVS and TPAVA can reduce more occurrences of hotspots by 47.13% and improve throughput by 6.50% on average. In particular, for four-layer 3D-MCPs, our on-line allocation strategy with VGVS and TPAVA averagely outperforms the STSC task scheduler by 10.97% on throughput.

$$HRR = \frac{\#hotspots(previous) - \#hotspots(proposed)}{\#hotspots(previous)} \quad (3)$$

$$TIR = \frac{throughput(proposed) - throughput(previous)}{throughput(previous)} \quad (4)$$

TABLE I
COMPARISON BETWEEN STSC [8] AND PROPOSED TASK SCHEDULER

# layers on 3D MCPs	Hotspot reduction (%)			Throughput improvement (%)		
	TPAVA	VGVS	TPAVA+ VGVS	TPAVA	VGVS	TPAVA+ VGVS
2layers	43.11	36.48	67.09	1.08	3.14	3.93
3layers	22.09	23.90	35.15	1.32	2.48	4.60
4layers	15.40	30.91	39.14	6.40	5.14	10.97
average	26.87	30.43	47.13	2.93	3.59	6.50

Furthermore, Table II shows the comparison in terms of hotspot reduction rate and throughput improvement rate between the proposed task scheduler and the Adapt3D task scheduler [11] under thermal constraints. Compared to the Adapt3D task scheduler, our on-line allocation strategy with VGVS and TPAVA can effectively reduce more hotspots by 53.91% and improve throughput by 32.06% on average. Moreover, as the scale of 3D MCPs increases, the improvement of throughput by TPAVA and VGVS can achieve 49.17% on average. As a result, the throughput by our on-line allocation strategy with TPAVA and VGVS is better than the STSC and the Adapt3D task schedulers, especially for larger-scale 3D-MCPs.

TABLE II
COMPARISON BETWEEN ADAPT3D [11] AND PROPOSED TASK SCHEDULER

# layers on 3D MCPs	Hotspot reduction (%)			Throughput improvement (%)		
	TPAVA	VGVS	TPAVA+ VGVS	TPAVA	VGVS	TPAVA+ VGVS
2layers	37.88	30.64	64.07	11.42	13.19	13.85
3layers	32.20	33.77	43.57	30.99	31.77	33.16
4layers	36.18	47.88	54.09	46.99	46.35	49.17
average	35.42	37.43	53.91	29.80	30.44	32.06

V. CONCLUSIONS

In this paper, a new thermal-constrained task scheduler based on a thermal-pattern-aware voltage assignment (TPAVA) for 3D-MCPs is proposed. It includes an on-line allocation strategy and a vertical-grouping voltage-scaling (VGVS) strategy to maximize the throughput and minimize the occurrences of hotspots. Rather than focusing on the thermal-aware allocation like previous works, this work targets on voltage assignment and voltage scaling instead. Through the analysis of temperature profiles, TPAVA assigns different initial operating-voltage levels to all cores for reducing the temperature increase in 3D-MCPs. Furthermore, VGVS triggers DVFS on all vertically-aligned cores of the overheated core for cooling overheated 3D-MCPs more quickly.

Experiments show that, compared to two previous thermal-constrained task schedulers (i.e. STSC [8] and Adapt3D [11]), the proposed task scheduler can reduce occurrences of hotspots by 47.13% and 53.91%, and improve throughput by 6.50% and 32.06%, respectively. As a result, the proposed task scheduler with TPAVA and VGVS significantly improves upon previous work by simultaneously reducing hotspots and optimizing throughput for 3D-MCPs under thermal constraints.

REFERENCES

[1] Y. Xie and Y. Ma, "Design space exploration for 3D integrated circuits," in *9th Int. Conf. on Solid-State and Integrated-Circuit Technology*, Beijing, 2008, pp. 2317–2320.

[2] K. Puttaswamy and G. H. Loh, "Thermal analysis of a 3D die-stacked high-performance microprocessor," in *ACM Great Lakes Symp. on VLSI, GLSVLSI*, 2006, pp. 19–24.

[3] W. L. Hung, G. Link, Y. Xie, N. Vijaykrishnan, and M. Irwin, "Interconnect and thermal-aware floorplanning for 3D microprocessors," in *ISQED*, March 2006, pp. 98–104.

[4] R. Viswanath, V. Wakharkar, A. Watwe, and V. Lebonheur, "Thermal performance challenges from silicon to systems," *Intel Technol. J.*, vol. 23, p. 16, 2000.

[5] M. Bakir, C. King, D. Sekar, H. Thacker, B. Dang, G. Huang, A. Naeemi, and J. Meindl, "3D heterogeneous integrated systems: Liquid cooling, power delivery, and implementation," in *CICC*, Sept. 2008, pp. 663–670.

[6] J. Cong, J. Wei, and Y. Zhang, "A thermal-driven floorplanning algorithm for 3D ICs," in *ICCAD*, Nov. 2004, pp. 306–313.

[7] J. Cong and Y. Zhang, "Thermal via planning for 3-d ics," in *ICCAD*, Los Alamitos, CA, USA, 2005, pp. 745–752.

[8] X. Zhou, J. Yang, Y. Xu, Y. Zhang, and J. Zhao, "Thermal-aware task scheduling for 3D multicore processors," *TPDS*, vol. 21, no. 1, pp. 60–71, Jan. 2010.

[9] A. Coskun, T. Rosing, K. Whisnant, and K. Gross, "Static and dynamic temperature-aware scheduling for multiprocessor SoCs," *TVLSI*, vol. 16, no. 9, pp. 1127–1140, sept. 2008.

[10] I. Yeo and E. J. Kim, "Temperature-aware scheduler based on thermal behavior grouping in multicore systems," in *DATE*, April 2009, pp. 946–951.

[11] A. Coskun, J. Ayala, D. Atienza, T. Rosing, and Y. Leblebici, "Dynamic thermal management in 3D multicore architectures," in *DATE*, 2009, pp. 1410–1415.

[12] C.-L. Lung, Y.-L. Ho, D.-M. Kwai, and S.-C. Chang, "Thermal-aware on-line task allocation for 3D multi-core processor throughput optimization," in *DATE*, march 2011, pp. 1–6.

[13] S. Murali, A. Mutapcic, D. Atienza, R. Gupta, S. Boyd, L. Benini, and G. De Micheli, "Temperature control of high-performance multi-core platforms using convex optimization," in *DATE*, 2008, pp. 110–115.

[14] V. Hanumaiah, R. Rao, S. Vrudhula, and K. Chatha, "Throughput optimal task allocation under thermal constraints for multi-core processors," in *DAC*, july 2009, pp. 776–781.

[15] "Intel 64 and IA-32 architecture software developer's manual,," http://support.intel.com/design/processor/manuals/.

[16] W. Huang, S. Ghosh, S. Velusamy, K. Sankaranarayanan, K. Skadron, and M. Stan, "Hotspot: a compact thermal modeling methodology for early-stage VLSI design," *TVLSI*, vol. 14, no. 5, pp. 501–513, may 2006.

[17] "HotSpot5.0 temperature modeling tool," http://lava.cs.virginia.edu/HotSpot/.

[18] A. Krum, *Thermal management*. The CRC handbook of thermal enginnering, F. Kreith ed., 2000, ch. 2.

[19] S. Wang and R. Bettati, "Reactive speed control in temperature-constrained real-time systems," in *ECRTS*, 2006, pp. 10–170.

[20] "Alpha 21264/ev6 microprocessor hardware reference manual," http://download.majix.org/dec/21264ev6_hrm.pdf.

[21] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan, "Temperature-aware microarchitecture: Modeling and implementation," *TACO*, vol. 1, no. 1, pp. 94–125, Mar. 2004.