Thermal Isolation Using Air Gap and Mechanically Flexible Interconnects for Heterogeneous 3-D ICs

Yue Zhang, Yang Zhang, Thomas Sarvey, Chaoqi Zhang, Muneeb Zia, and Muhannad Bakir

***Abstract*— This paper presents a set of thermal isolation technologies to provide a measure to thermally insulate low-power temperature-sensitive tiers from the time-varying power dissipation of high-power tier in heterogeneous 3-D integrated circuits. The proposed technologies use an air gap and mechanically flexible interconnects (MFIs) to replace conventional microbumps and underfill. A two-tier testbed is fabricated to emulate a heterogeneous 3-D stack. The results are then benchmarked with a standard 3-D stacking approach that is simulated using finite-volume modeling. Compared with the conventional approach using microbumps and underfill, a temperature reduction of 35.9% can be achieved in the low-power tier by implementing the air gap and MFIs. Four-point and daisy-chain resistances of the MFIs are measured to verify the electrical connectivity between the tiers during temperature cycling.**

***Index Terms*— 3-D integrated circuits (ICs), heterogeneous integration, thermal isolation.**

I. Introduction

A KEY challenge for 3-D integrated circuits (ICs) is thermal management. There are two main thermal challenges in typical 3-D ICs. First, in a homogeneous integration of multiple high-power tiers, a cooling solution that scales with the number of dice in the stack is needed [1]. Second, in heterogeneous integration, a thermal isolation solution may be necessary to protect a low-power tier from a high-power tier [2], [3]. While substantial research studies have focused on heat removal within homogeneous integration with multiple high-power tiers [4]–[7], very little effort has investigated thermal isolation of low-power tiers from high-power tiers in a 3-D stack.

In today's approach to 3-D IC stacks, tiers are bonded using microbumps along with underfill, which is applied between tiers to alleviate the thermomechanical stress on the solder microbumps. However, the thermal conductivity of underfill is approximately 0.4–1.3 W/mK. This will introduce a small thermal resistance between the two tiers and cause thermal coupling between the tiers. The thermal coupling challenge is especially significant in heterogeneous integration where high-power tiers are stacked with low-power tiers, such as a memory-on-logic, a logic-on-silicon photonic, and Microelectromechanical systems-on-logic chip stacks. Without effective thermal isolation between the tiers, the thermal coupling will cause the low-power and temperature-sensitive tier to follow the temperature profile of the high-power tier and, thus, leading to possible performance degradation. For example, stacking a silicon photonic chip in a stack adjacent to logic and memory has been explored [8], [9]. However, some key components in the photonic chip, such as microring modulators, are sensitive to temperature fluctuations. A local heater is often used to create a constant temperature environment in a local region. Extra tuning power is needed for the heater if the generated heat spreads to the adjacent area because of thermal coupling. In a second example, it has been shown that the temperature of SRAM increases by 30 °C–40 °C [10] under the influence of adjacent logic die, leading to undesirable leakage power increase and performance degradation. Finally, in a DRAM-on-logic stack, investigated in [11], the temperature of the DRAM die is close to that of the logic although the DRAM die has much smaller power dissipation, which indicates strong thermal coupling between the two tiers.

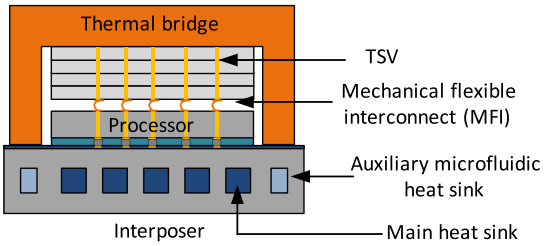


Fig. 1. Illustration of a 3-D IC application of memory-on-processor with the proposed thermal isolation technologies to replace microbumps and underfill with air gap and thermally degraded MFIs.

Integrating an air gap or vacuum between tiers in 3-D ICs is proposed as a solution to protect the low-power tier [3], [12]. It is shown that by integrating an air-gap layer between the memory and the logic, the temperature of the memory tier can be reduced by 25.7 °C [3]. However, the proposed approach in [3] did not address thermomechanical challenges nor reported experimental results of the thermal isolation approach. To this end, we propose to integrate an air gap and mechanically flexible interconnects (MFIs) to replace both the microbumps and the underfill. Fig. 1 is an application example of memory-on-processor where the proposed thermal isolation technologies can be used. Unlike rigid solder microbumps,

Manuscript received May 21, 2015; revised September 23, 2015 and November 10, 2015; accepted November 13, 2015. This work was supported by the Defense Advanced Research Projects Agency under Grant N66001-12-1-4240. Recommended for publication by Associate Editor P. Franzon upon evaluation of reviewers' comments.

The authors are with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: yzhang324@gatech.edu; steven.zhang@gatech.edu; tsarvey@gatech.edu; chqzhang@gatech.edu; muneeb.zia@gatech.edu; muhannad.bakir@mirc.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCPMT.2015.2505670

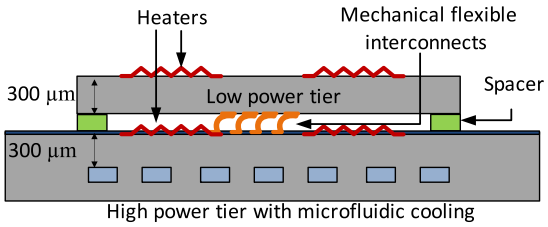
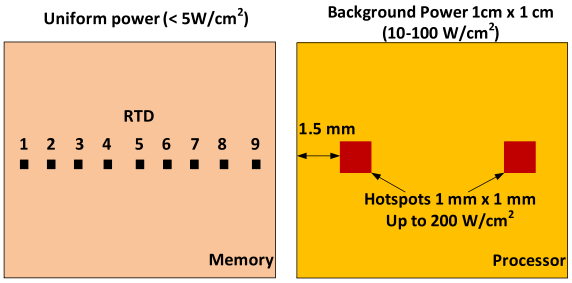


Fig. 2. Schematic of the designed testbed for the evaluation of the proposed thermal isolation technologies.



(a) (b)

Fig. 3. Schematics of the power map designs of (a) top tier (low-power tier) and (b) bottom tier (high-power tier).

MFIs can deform elastically under stress, which helps maintain the electrical connectivity between the memory stack and the processor [13]. Due to this behavior, MFIs can help eliminate the underfill and, thus, reduce the thermal coupling between tiers. Once the memory stack is effectively thermally isolated from the processor, it is also isolated from a cooling path through the bottom. Thus, a thermal bridge can be attached to the top of the stack to provide an alternative cooling path, as previously discussed in [3]. With the combined thermal isolation technologies and the thermal bridge, new opportunities for improved heterogeneous system integration and miniaturization become possible.

The goal of this paper is to demonstrate the thermal isolation technologies using the MFIs and the air gap. A simplified two-tier stack testbed, as shown in Fig. 2, is designed, fabricated, and tested in this paper to emulate a stack of low-power and high-power tiers. The experimental results are then benchmarked with a standard 3-D stacking scenario that is modeled by a finite-volume modeling method [3]. The model is used to simulate a two-tier stack with both thermal isolation technologies and the thermal bridge.

II. Design of the Testbed

Guided by the previous modeling and analysis [3], a thermal testbed is designed, as shown in Fig. 2, to explore thermal coupling and solutions. The MFIs are designed to be clustered in the middle region to further enhance the thermal isolation [3].

The designed and fabricated testbed consists of low-power and high-power tiers to emulate a heterogeneous 3-D stack. The testbed is designed to emulate a memory-on-processor application although the technologies apply to other stacks including photonics. The microfluidic heat sink (MFHS) is integrated in the high-power tier (bottom tier). MFIs are used as interconnects between the two tiers (instead of microbumps). Fig. 3(a) shows the power map and temperature sensor designs for the low-power tier. The low-power tier

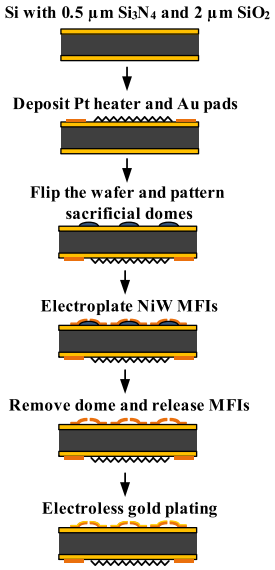


Fig. 4. Process flow for the low-power tier.

dissipates a uniform power of less than 5 W. A spiral heater is formed over a 1-cm × 1-cm area. Nine resistance temperature detectors (RTDs) are inserted along the middle of the chip in order to measure temperature along the length of the chip (in other words, these temperature sensors are designed to capture the temperature gradient across the die). Each RTD has dimensions of 500 μm × 88 μm and yields a resistance of ∼200 Ω. Since the MFIs are clustered in the middle region, the thermal coupling between the tiers is expected to be nonuniform across the chip, in particular from the center to the edges. The thermal bridge is not included in the testbed. The cooling of the top tier is thus through natural convection, which limits the power that can be applied to the low-power tier.

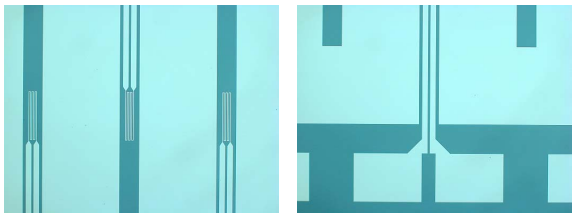
Fig. 3(b) shows a schematic of the high-power tier. The chip area is 1 cm × 1 cm. There are two hotspots on the chip each measuring 1 mm × 1 mm. The hotspots are located 1.5 mm away from the edges.

The two chips are interconnected with an array of gold-passivated NiW MFIs. NiW is chosen as the core material for the MFIs because of its high yield strength (up to 1.9 GPa) and the CMOS-compatible fabrication approach through electroplating [13]. The array contains 12 × 100 MFIs yielding a total of 1200 MFIs. This number is chosen based on the wide I/O specifications [14]. The MFI design has a pitch of 75 μm × 100 μm. The entire MFI array is 9940 μm × 870 μm. Four-point and daisy-chain resistance measurements of 38 MFIs are designed and integrated into the layouts to enable postassembly electrical resistance verification.

III. Testbed Fabrication and Test Setup

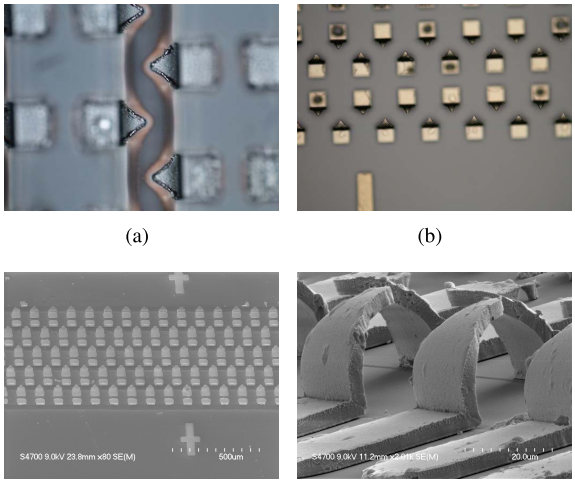
A. Testbed Fabrication

The process flows for the two tiers (low-power and high-power tiers) are discussed in this section. For the low-power tier, the process begins with a double side-polished, 300-μm-thick Si wafer (Fig. 4). The bottom side of the wafer has 0.2-μm-thick Si_3N_4 for MFIs formation, and the



(a) (b)

Fig. 5. Images of (a) parts of the RTD array and (b) pad of the RTD.



(a) (b) (c) (d)

Fig. 6. Images of (a) MFIs electroplated on top of the polymer dome, (b) free standing MFIs after dome removal, (c) SEM of the MFI array from top, and (d) side view of the MFIs.

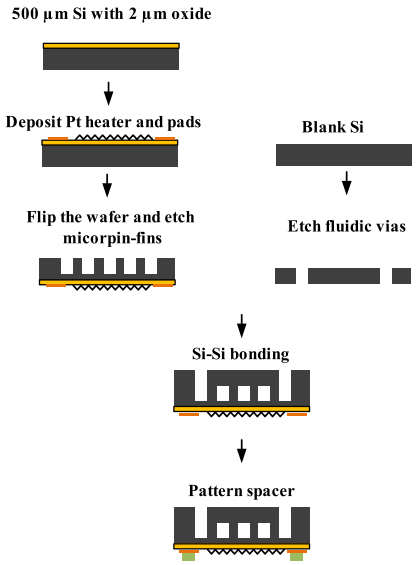
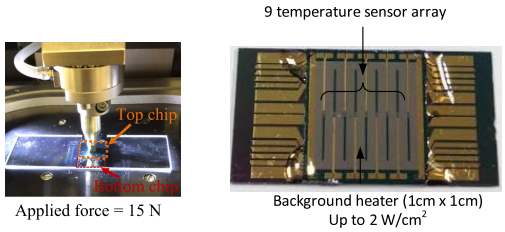


Fig. 7. Process flow for the high-power tier.



(a) (b)

Fig. 8. (a) Flip chip bonding assembly and (b) assembled two-tier testbed.

top side has 2-μm-thick SiO_2 for heater and temperature sensor formation. Next, 0.2-μm-thick Pt-based heaters/RTD are formed on the top side using liftoff. Fig. 5 shows the nine RTDs and the RTD pads on the top die. The next step is to deposit 0.5-μm-thick gold pads above the Pt RTD contact pads. The gold pads facilitate wire bonding, which is needed during testbed assembly. The sample is next flipped over for MFI-related processes. SPR220 is spun and patterned to form sacrificial squares. The squares then undergo a reflow process to form a dome structure with a height of 20 μm [15]. The wafer is then placed in an NiW electroplating solution to electroplate the MFIs to a thickness of 4.5 μm. After removing the sacrificial polymer dome beneath the MFIs, the MFIs become freestanding. Fig. 6(a) shows the NiW MFIs electroplated above the sacrificial dome. The last fabrication step for the low-power tier is to passivate the MFI surface with gold using electroless plating. The gold passivation prevents NiW from oxidizing and also provides a lower electrical contact resistance, as shown in Fig. 6(b) [13]. Fig. 6(c) and (d) shows the SEM images of the freestanding MFIs. The final height of the MFIs is 25 μm.

The process steps involved in the fabrication of the high-power tier are shown in Fig. 7. The process starts with a double side-polished 500-μm-thick Si wafer. Since the micropin-fins are integrated into this tier, the wafer is chosen to be thicker to provide enough mechanical stability. The next step involves the deposition of a 2-μm-thick SiO_2 layer on the top side. Next, 0.2-μm-thick Pt heaters/RTDs and 0.5-μm-thick gold pads are patterned using two liftoff steps. The wafer is next flipped over and 200-μm-deep micropin-fins are etched using a standard Bosch etching process. Fluidic vias are etched on a second wafer that is 300-μm thick and serves as a cover. The two wafers are then bonded using Si–Si fusion bonding and undergo an annealing process at 400 °C. The final step for preparing the high-power tier is to deposit polymer pillars on the heater side. The pillars serve as spacers to ensure a gap of greater than 10 μm.

B. Assembly

The wafer is next diced. The low-power and high-power tiers are then assembled using a Finetech submicrometer flip-chip bonder, as shown in Fig. 8(a). After aligning the two tiers, the alignment head is placed in contact with the stack and a force of 15 N is applied to bond the stack; while under compression, Devcon 5 Minute epoxy is applied to the four corners to hold the tiers in position. The force is released once the epoxy hardens. Fig. 8(b) is an image of the assembled testbed. Fig. 9(a) shows an X-ray image of the bonded sample. The region within the black square is magnified and shown in Fig. 9(b). The micropin-fins, the four-point resistance measurement structures, and the daisy-chain resistance measurement structures can be seen in Fig. 9(b).

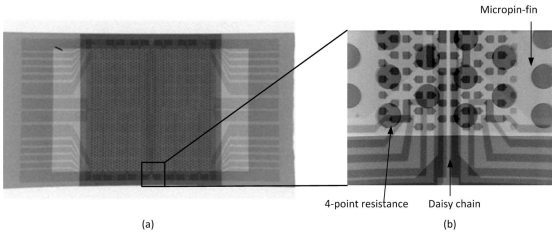


Fig. 9. X-ray of (a) overall view of the bonded chip and (b) magnified view.

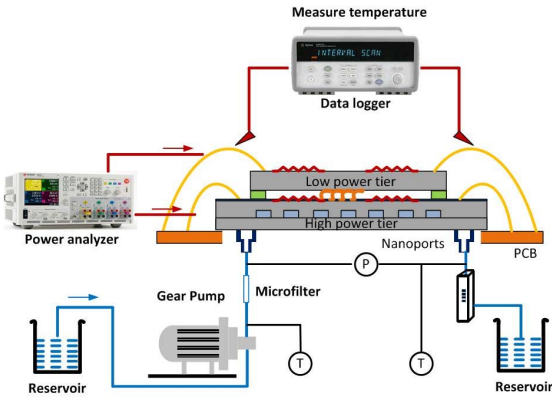


Fig. 10. Microfluidic test setup to evaluate the thermal isolation technologies.

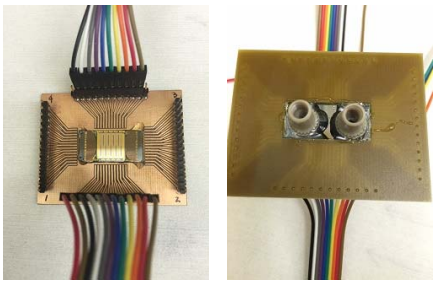


Fig. 11. (a) Top and (b) bottom view of the stack assembled to a PCB board using wire bonding.

C. Thermal and Electrical Test Setup

The microfluidic test setup is shown in Fig. 10. Given the micropin-fins are only etched in the high-power tier (bottom tier), the coolant only flows within the high-power tier. The top tier is bonded to the bottom tier through MFIs that are located in the center region. The stack is then bonded to a predesigned Printed circuit board (PCB) for testing [Fig. 11(a)]. Nanoports (i.e., fluidic ports) are attached on the bottom of the sample, as shown in Fig. 11(b). An Agilent data logger is used to source current into the on-chip heater/RTDs on both tiers. The data logger is used to measure the resistance of the RTDs on the top and bottom tiers and extract the junction temperatures [6].

Fig. 12 shows the test setup for the four-point resistance measurement of a single MFI. The measured resistance consists of the electrical resistance of a single MFI and contact resistance between the MFI and the gold pad.

IV. Thermal and Electrical Experimental Results

To demonstrate the thermal isolation concept, several test cases are emulated. In all tests, the inlet Deionized water water

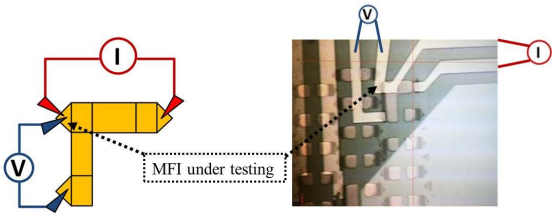


Fig. 12. Four-point resistance measurement of MFI.

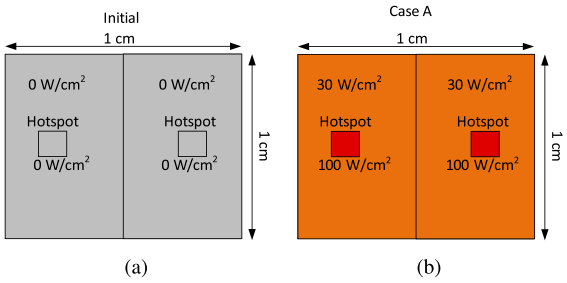


Fig. 13. (a) Initial case when the high-power tier dissipates 0 W. (b) Case A where the background power density is 30 W/cm^2 and the hotspot power density is 100 W/cm^2.

temperature is 19.5 °C ± 0.5 °C. The room temperature is 22.5 °C ± 0.5 °C. In Sections IV-A, IV-B and IV-C, the test cases are compared in pairs to better understand the impact of thermal isolation. The junction temperature at the center of the chip is computed as the average of the left and right background temperatures.

A. Thermal Testing I: Powering the High-Power Tier

In a heterogeneous 3-D stack consisting of a low-power die bonded on top of a high-power die using microbumps and underfill, thermal coupling is expected to result in a temperature increase in the low-power die when the high-power tier is powered. In this section, this scenario is emulated using the thermal isolation testbed.

The power maps of the high-power tier are shown in Fig. 13. The low-power tier dissipates 0.5 W in all the evaluated cases. In the initial case, the high-power tier does not dissipate any power. In Case A, the background power density is 30 W/cm^2, while the two hotspots dissipate 100 W/cm^2. The junction temperature across the two tiers is shown in Fig. 14. In the initial case, the temperature of both tiers is close to the inlet water temperature. When the bottom tier is powered to a background power density of 30 W/cm^2 and a hotspot power density of 100 W/cm^2 [power map shown in Fig. 13(b)], the temperature of the bottom tier increases at all locations. The temperature of the left and right side background heaters increases to 28.6 °C and 32.4 °C, respectively. The temperature of the left and right side hotspots increases to 31.6 °C and 36.4 °C, respectively. However, the average temperature of the upper tier increases to 26.9 °C. The temperature of the upper tier follows the same temperature trend of the bottom tier. However, owing to the thermal isolation technology, the temperature increase is not as high as the bottom tier. An interesting point to be noted is that the highest temperature

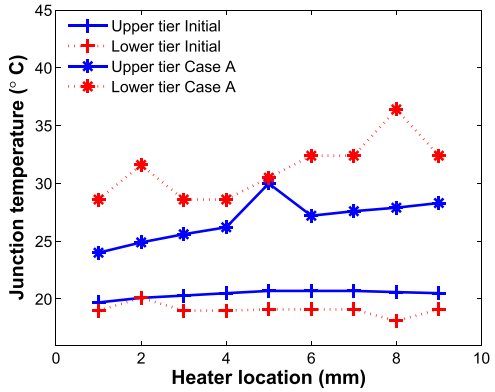


Fig. 14. Measured junction temperature fluctuation before and after the high-power tier is powered.

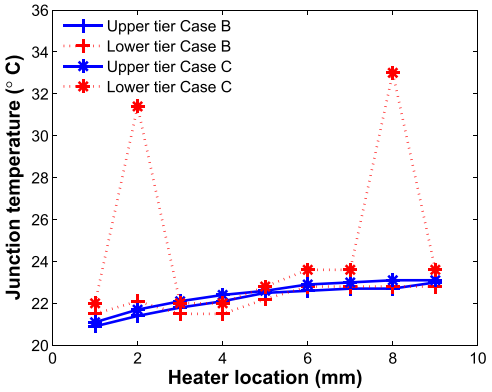


Fig. 16. Measured junction temperature fluctuation of top and bottom tiers in Case B and Case C.

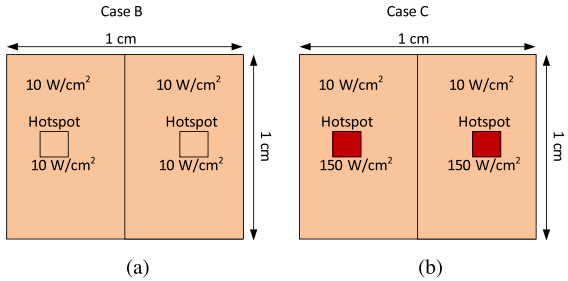


Fig. 15. (a) Uniform power density of 10 W/cm^2 in the bottom tier (Case B). (b) Background power of 10 W/cm^2 plus two hotspots each dissipating 150 W/cm^2 (Case C).

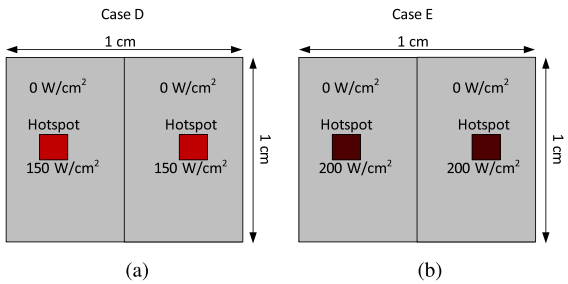


Fig. 17. (a) Zero background power with two hotspots each dissipating 150 W/cm^2 (Case D). (b) Zero background power with two hotspots each dissipating 200 W/cm^2 (Case E).

in the upper tier is located at the center of the die. Moreover, the temperature of both tiers is very close at the center. This effect can be attributed to the dense MFI array that is clustered in the middle and, thus, creates an enhanced thermal path. Another point to be noted is that the temperature of the upper die gradually increases from inlet to outlet. One reason is that the top tier follows the same temperature trend of the bottom tier. The other reason is due to the imperfection of the testbed. In the testbed, epoxy is used at the four corners to securely bond the upper die to the bottom die, and thus, heat can be conducted through the epoxy. When the temperature of the coolant becomes elevated at the outlet, it impacts the temperature of the upper die at the outlet. Therefore, the temperature at location 9 [refer to Fig. 3(a)] is higher than that at location 1. The epoxy thermal coupling is considered in the finite-volume modeling in the following sections.

B. Thermal Testing II: Minimizing Hotspot Coupling

Hotspot cooling is a critical issue for today's high-performance computers. By stacking a low-power die with a high-power die, the hotspots can also occur in the low-power die because of thermal coupling. This presents a number of challenges for temperature-sensitive low-power dice. Therefore, in this section, we explore the effects of hotspots further.

The power maps of the simulated cases are shown in Fig. 15. In Fig. 15(a), the bottom tier dissipates 10 W/cm^2 across the chip. The junction temperature for each location on both tiers is shown in Fig. 16 (Case B). Next (Case C), the power density of the two hotspots increases to 150 W/cm^2, while the background power remains unchanged [Fig. 15(b)]. The corresponding temperature of each chip is shown in Fig. 16 (Case C). In Case B, the temperature curve is relatively flat indicating uniform temperature without hotspots. When the power density of the hotspot increases, one obvious observation is that there are two peak temperatures that occur in the bottom die. This is expected because of the large power density of the hotspots. The two peak temperatures are 31.4 °C and 33 °C, respectively. However, also in Case C, there are no obvious hotspots in the upper tier. The temperature of the upper tier gradually increases from 21.1 °C to 23.1 °C. This demonstrates that the proposed thermal isolation concept effectively minimizes the hotspot coupling between the stacked tiers.

To illustrate this point, two extreme cases are emulated where only the hotspot regions are powered while the background dissipates no power. In addition, the power density of the hotspot increases to 200 W/cm^2. The two power maps are shown in Fig. 17(a) and (b). The corresponding temperature in the two cases is shown in Fig. 18. In these two cases, the temperature of the bottom tier is close to room temperature except for the two hotspots, where the temperatures are 29.3 °C and 33 °C for Case D and Case E, respectively. Even though the power density of the hotspot is high, the total power is low, since each hotspot is only 1 mm × 1 mm. The fluid temperature barely increases after flowing through the hotspot. Thus, the hotspots near the inlet and the outlet have the same temperature. In Cases D and E, the temperature at location 2 in the upper tier is 20 °C and 20.2 °C, respectively. However, the temperature at location 2

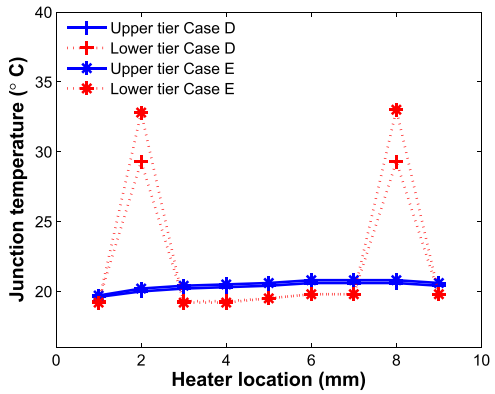


Fig. 18. Measured junction temperature fluctuation of the top and bottom tiers in Case D and Case E.

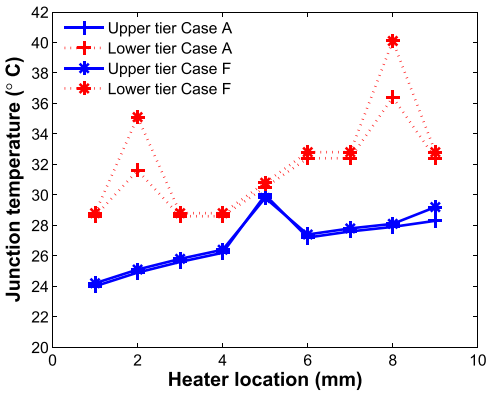


Fig. 20. Measured junction temperature fluctuation of the top and bottom tiers in Case A and Case F.

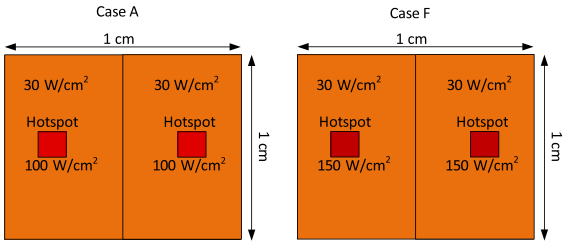


Fig. 19. (a) Background power of 30 W/cm^2 plus two hotspots each dissipating 100 W/cm^2 (Case A). (b) Background power of 30 W/cm^2 plus two hotspots each dissipating 150 W/cm^2 (Case F).

in the bottom tier is 29.3 °C and 32.8 °C, respectively. The maximum junction temperature difference is 12.6 °C. For reference, in Fig. 14, when the bottom chip is not dissipating any power, the junction temperature at location 2 in the upper tier is also 20 °C. The temperature barely changes in the upper tier after the hotspot power increases. This demonstrates that thermal isolation using MFIs has greatly decreased the hotspot coupling between tiers.

C. Thermal Testing III: Increasing the Power of the Bottom Tier

In Section IV-B, we demonstrated that the thermal isolation technology under consideration can prevent vertical coupling and, thus, protect the low-power tier from the hotspots in the high-power tier. In this section, the bottom tier dissipates an elevated power density of 30 W/cm^2 in addition to the hotspots. In the two cases, the two hotspots dissipate 100 and 150 W/cm^2, respectively, as shown in Fig. 19. The corresponding temperature of the two tiers in the two cases is shown in Fig. 20.

At location 2 and location 8, where the hotspots are located, the temperature difference between the top and bottom tiers is large. For example, the temperature difference between the two tiers at location 8 in Case F is 12 °C. However, at other locations, the temperature difference is less (relative to hotspot locations). For example, the temperature at location 9 in the upper and bottom tiers is 32.8 °C and 29.2 °C, respectively. The thermal isolation effect is weakened in this case, since the coolant temperature increases as it absorbs heat from the bottom tier. The higher the power the bottom tier dissipates, the greater is coolant temperature. The elevated coolant temperature in turn causes the temperature to increase in the upper tier. This effect is expected when the two tiers share the same MFHS especially for the locations near the outlet. On the other hand, the measured temperature in the bottom tier is an average of half of the die. The actual temperature of location 9 should be higher than the average temperature. One method to minimize the impact of coolant temperature increase in the upper tier is to allocate an independent MFHS for it, as shown in Fig. 1.

Another interesting observation is that the temperature is highest at the center of the upper tier and similar to that of the bottom tier. The reason is that the MFIs are densely clustered in the middle, which complies with wide I/O technology. In wide I/O technology, all Through silicon vias are located in a rectangular array in the middle of the chip.

In our designed two-tier testbed, the hotspots always reside away from the center of the bottom chip. In applications where the hotspots appear in the middle of the bottom chip, the thermal coupling will increase using the current MFI layout, in which case redesigning of the MFI layout will become necessary.

D. Electrical Testing of MFIs

In order to demonstrate the electrical connectivity between the two tiers after bonding, two sets of electrical resistance measurements are performed. The four-point resistance measurements are made at four different locations on the sample. The measured average electrical resistance is 46.5 mΩ. The measured resistance consists of the resistance of the MFI, the part of the landing pad, and the contact resistance. The daisy-chain resistance of 38 MFIs is also measured during the thermal measurements. At room temperature, the resistance of the daisy chain including the leading wires is 19.55 Ω. When the temperature of the bottom chip increases, the daisy-chain resistance also increases. The highest measured resistance during all thermal testing is 19.77 Ω. The daisy-chain resistance measurement results provide confidence that all the electrical contacts through the MFIs remain throughout thermal testing.

TABLE I
PARAMETERS USED IN THE FINITE-VOLUME MODEL

	Conductivity (W/mK)	Thickness (μm)
Memory die	149	300
Underfill layer	3	25
Air gap	2.4E-2	25
Processor die	149	300
Micro-bump	60	25
Copper	400	N/A
MFIs	10	25
SiO_2	1.38	2.5

TABLE II
BOUNDARY CONDITIONS ASSUMED IN THE FINITE-VOLUME MODEL

Stack	Boundary face	Heat Transfer Coefficient (W/Km^2)
Conventional microbump stack	Top	10
	MFHS (bottom)	52000
	Others(near adiabatic)	5
Our proposed stack	Top	13000
	MFHS (bottom)	52000
	Side (near adiabatic)	5

V. BENCHMARK WITH CONVENTIONAL 3-D STACKING APPROACH WITH MICROBUMPS AND UNDERFILL

A. Validation by Finite-Volume Modeling

A finite-volume modeling is used to validate our experimental results. The assumptions of the boundary conditions used in the model are made based on the measurement results. For microfluidic cooling in the bottom tier, the convective heat transfer coefficient is assumed to be 5.2×10^4 W/m^{2}K. For the top tier, a heat transfer coefficient of 1.3×10^4 W/m^{2}K is assumed. This convective boundary condition is only applied on the edges of the top tier. The calculation is based on the initial case [Fig. 13(a)]; the top tier dissipates 0.5 W, while the bottom tier dissipates 0 W. The equivalent thermal resistance is calculated to be 3.4 K/W for the top chip. The MFIs are represented by a 1-cm $\times$ 1-mm rectangular layer with an equivalent thermal conductivity. All the parameters and boundary conditions used in the model are included in Tables I and II. The measured and modeled results for Case F are shown and compared in Fig. 21. In all the simulated cases, the error is within 2 °C.

B. Comparison With Microbump and Underfill Using Finite-Volume Modeling

To benchmark with conventional 3-D integration scenarios, the same stack is modeled with microbumps and underfill, as shown in Fig. 22(b) and (c). In Fig. 22(b), the microbumps are uniformly distributed, while in Fig. 22(c), they are clustered in the middle as specified in the wide I/O technology. For a fair comparison, the same number of MFIs and microbumps is assumed. Tables I and II list the parameters used in the model. The power map in Case F is used in this simulation. In this case, the background power density is 30 W/cm^2, while the hotspots dissipate 150 W/cm^2. For the two cases without thermal isolation, as shown in Fig. 23(b) and (c), there is little difference between the two cases because the heat conduction through underfill dominates. In Fig. 23(b) and (c), we can see that the temperature of the upper tier follows that of the lower tier. In most locations, the temperature is similar in both tiers.

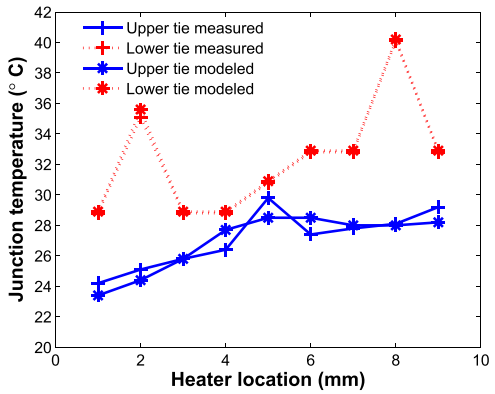


Fig. 21. Comparison between the measured and modeled junction temperatures in Case F.

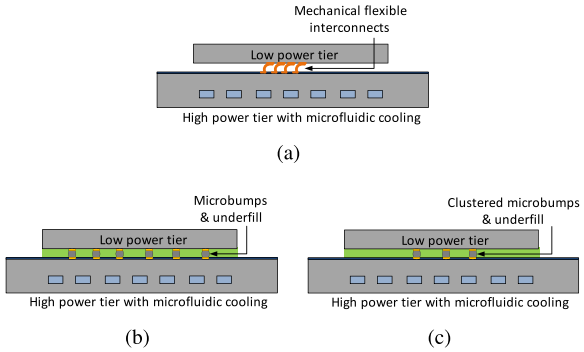


Fig. 22. Modeled heterogeneous stack with (a) MFI and air cavity, (b) uniform microbumps and underfill, and (c) clustered microbumps and underfill.

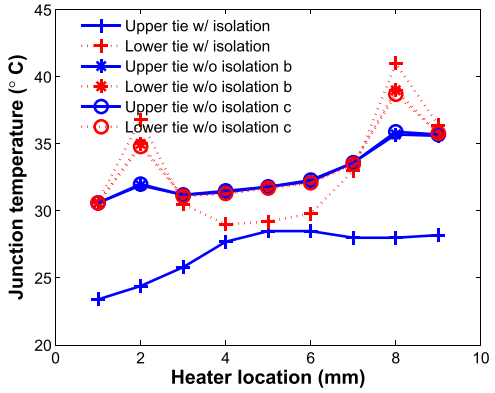


Fig. 23. Modeled chip temperature in both tiers with and without the thermal isolation. In the case without thermal isolation, microbumps and underfill are integrated between the tiers.

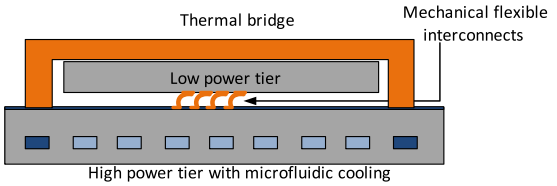


Fig. 24. Heterogeneous 3-D stack with MFIs and independent MFHS for the low-power die.

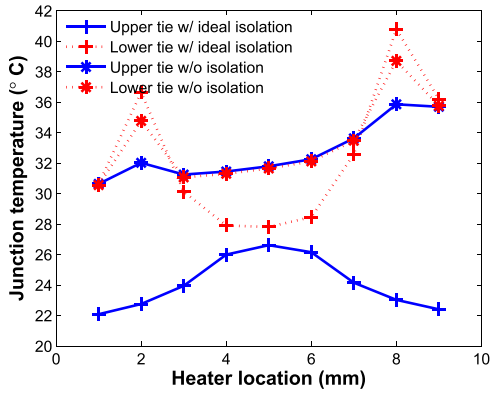


Fig. 25. Benchmark the ideal thermal isolation technology with the conventional 3-D stacking approach using finite-volume modeling.

At the hotspot near the outlet, the temperature of the upper tier and the lower tier is 35.9 °C and 38.7 °C, respectively. For the case with thermal isolation, the temperature of the upper tier and the lower tier is 28 °C and 41 °C, respectively. The thermal isolation technology is shown to reduce the upper-tier temperature by 8 °C at location 8 and, thus, yields a temperature reduction of 19.5%, while the bottom-tier temperature increases by 2 °C. This is because the upper tier helps to spread the heat and, thus, lower the hotspot temperature of the bottom tier. At locations without hotspots, such as location 9, thermal isolation is also observed. In the case without thermal isolation, the temperature of the upper tier and the lower tier is 35.7 °C and 35.8 °C, respectively, while with thermal isolation, the temperature of the upper tier drops to 28.2 °C without causing the temperature of the lower tier to increase.

C. Thermal Isolation With Independent MFHS Dedicated to the Low-Power Tier

The implication from the analysis in the previous section is that allocating an independent MFHS to the low-power tier may further decouple it from the high-power die. Therefore, the proposed concept with thermal bridge and independent MFHS is modeled (as shown in Fig. 24) and benchmarked with the conventional microbump and underfill approach. In the ideal thermal isolation case, the temperature of the high-power and low-power tiers at location 8 is 40.8 °C and 23 °C, respectively (Fig. 25). While in the conventional bonding scenario, as shown in Fig. 22(b), the temperature of the high-power and low-power tiers is 38.7 °C and 35.9 °C, respectively. A temperature reduction of 35.9% is achieved in the low-power tier by using MFIs and independent MFHS.

VI. Conclusion

For heterogenous 3-D integration, including high-power and low-power dice (e.g., memory and nanophotonics), thermal coupling is a critical issue. The proposed thermal isolation technology features an air cavity between the tiers, MFIs as interconnects, and a thermal bridge for the isolated tiers. To demonstrate the thermal isolation technology, a two-tier testbed with heterogeneous elements is designed, fabricated, and tested. Various thermal test cases are evaluated. The proposed technology effectively decouples the two tiers thermally. One case shows that the proposed technology effectively prevents hotspots on the high-power chip from coupling to the low-power tier; a best case of 35.9% reduction in the low-power tier can be achieved. Four-point resistance measurements of the MFIs are performed along with MFI daisy-chain resistance measurements that confirm MFI electrical connectivity between tiers throughout all power cycling.

Acknowledgment

The authors would like to thank S. Hu for performing wire bonding on the testbed.

References

[1] W. R. Davis *et al.*, "Demystifying 3D ICs: The pros and cons of going vertical," *IEEE Des. Test Comput.*, vol. 22, no. 6, pp. 498–510, Nov./Dec. 2005.

[2] Y. Zhang, H. Oh, and M. S. Bakir, "Within-tier cooling and thermal isolation technologies for heterogeneous 3D ICs," in *Proc. IEEE Int. Conf. 3D Syst. Integr. (3DIC)*, Oct. 2013, pp. 1–6.

[3] Y. Zhang, Y. Zhang, and M. S. Bakir, "Thermal design and constraints for heterogeneous integrated chip stacks and isolation technology using air gap and thermal bridge," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 4, no. 12, pp. 1914–1924, Dec. 2014.

[4] D. B. Tuckerman and R. F. W. Pease, "High-performance heat sinking for VLSI," *IEEE Electron Device Lett.*, vol. 2, no. 5, pp. 126–129, May 1981.

[5] S. Ndao, Y. Peles, and M. K. Jensen, "Multi-objective thermal design optimization and comparative analysis of electronics cooling technologies," *Int. J. Heat Mass Transf.*, vol. 52, nos. 19–20, pp. 4317–4326, Sep. 2009.

[6] Y. Zhang, A. Dembla, and M. S. Bakir, "Silicon micropin-fin heat sink with integrated TSVs for 3-D ICs: Tradeoff analysis and experimental testing," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 3, no. 11, pp. 1842–1850, Nov. 2013.

[7] Y. Zhang *et al.*, "Coupled electrical and thermal 3D IC centric microfluidic heat sink design and technology," in *Proc. IEEE Electron. Compon. Technol. Conf. (ECTC)*, May/Jun. 2011, pp. 2037–2044.

[8] D. Brunina, D. Liu, and K. Bergman, "An energy-efficient optically connected memory module for hybrid packet- and circuit-switched optical networks," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar./Apr. 2013, Art. ID 3700407.

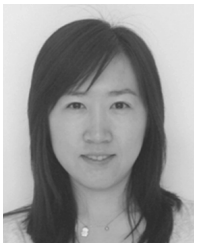
[9] Z. Li *et al.*, "Reliability modeling and management of nanophotonic on-chip networks," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 1, pp. 98–111, Jan. 2012.

[10] S. Chatterjee, M. Cho, R. Rao, and S. Mukhopadhyay, "Impact of die-to-die thermal coupling on the electrical characteristics of 3D stacked SRAM cache," in *Proc. IEEE Semiconductor Thermal Meas. Manage. Symp. (SEMI-THERM)*, Mar. 2012, pp. 14–19.

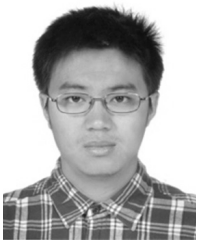
[11] H. Oprins, V. O. Cherman, B. Vandevelde, G. Van der Plas, P. Marchal, and E. Beyne, "Numerical and experimental characterization of the thermal behavior of a packaged DRAM-on-logic stack," in *Proc. IEEE 62nd Electron. Compon. Technol. Conf. (ECTC)*, May 2012, pp. 1081–1088.

[12] P. Franzon, J. Wilson, and M. Li, "Thermal isolation in 3D chip stacks using vacuum gaps and capacitive or inductive communications," in *Proc. IEEE Int. 3D Syst. Integr. Conf. (3DIC)*, Nov. 2010, pp. 1–4.

[13] C. Zhang, H. S. Yang, and M. S. Bakir, "Highly elastic gold passivated mechanically flexible interconnects," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 3, no. 10, pp. 1632–1639, Oct. 2013.
[14] *WIDE I/O SINGLE DATA RATE (WIDE I/O SDR)*, JEDEC Standard JESD229, Dec. 2011. [Online]. Available: http://www.jedec.org/standards-documents/docs/jesd229
[15] C. Zhang, H. S. Yang, and M. S. Bakir, "Mechanically flexible interconnects (MFIs) with highly scalable pitch," *J. Micromech. Microeng.*, vol. 24, no. 5, pp. 1–8, May 2014.



Yue Zhang (S'13) received the B.S. degree in electrical engineering and automation from the Harbin Institute of Technology, Harbin, China, and the University of Science and Technology of Lille, Lille, France, in 2007, and the M.S. degree in micro and nano technology from the University of Science and Technology of Lille, in 2009, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2015.



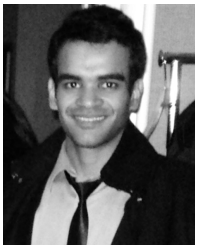
Yang Zhang (S'13) received the B.S. degree in microelectronics and math (double major) from Peking University, China, in 2012. He is currently pursuing the Ph.D. degree in electrical engineering with the Georgia Institute of Technology, Atlanta, GA.

Thomas Sarvey received the B.S. degrees in computer engineering and physics from the University of Maryland, College Park, in 2012. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Georgia Institute of Technology in Atlanta, Georgia, USA.

Chaoqi Zhang (S'12) received the B.S. degree in physics from Shandong University, Jinan, China, in 2005, the M.S. degree in microelectronics from Tsinghua University, Beijing, China, in 2008, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2015.

He is currently a Senior Hardware Engineer with the Oracle Networking Group, Oracle, San Diego, CA, where he is involved in advanced integration and packaging for silicon-photonic device and system.

Dr. Zhang was the recipient of the 2014 IEEE Transactions on Components, Packaging and Manufacturing Technology best paper award for advanced packaging and a reviewer for IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY and IEEE TRANSACTIONS ON ELECTRONICS DEVICE.



Muneeb Zia received the B.S. degree in electronic engineering from the GIK Institute of Engineering Sciences and Technology, Topi, Pakistan, in 2010, and the M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2013, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

His current research interests include fabrication and characterization of flexible interconnects for 2.5D and 3D systems, novel biosensing platforms and silicon photonics packaging.

Muhannad Bakir (SM'12) received the B.E.E. degree from Auburn University, Auburn, AL, in 1999 and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech) in 2000 and 2003, respectively.

He is currently an Associate Professor and the ON Semiconductor Junior Professor in the School of Electrical and Computer Engineering at Georgia Tech. His areas of interest include three-dimensional (3D) electronic system integration, advanced cooling and power delivery for 3D systems, biosensors and their integration with CMOS circuitry, and nanofabrication technology.

Dr. Bakir is the recipient of the 2013 Intel Early Career Faculty Honor Award, 2012 DARPA Young Faculty Award, 2011 IEEE CPMT Society Outstanding Young Engineer Award, and was an Invited Participant in the 2012 National Academy of Engineering Frontiers of Engineering Symposium. In 2015, he was elected by the IEEE CPMT Society to serve as a Distinguished Lecturer for a four-year term. He and his research group have received fourteen conference and student paper awards including five from the IEEE Electronic Components and Technology Conference (ECTC), four from the IEEE International Interconnect Technology Conference (IITC), and one from the IEEE Custom Integrated Circuits Conference (CICC). His group was awarded the 2014 Best Paper of the IEEE TRANSACTIONS ON COMPONENTS PACKAGING AND MANUFACTURING TECHNOLOGY in the area of advanced packaging. He is an Editor of IEEE TRANSACTIONS ON ELECTRON DEVICES and an Associate Editor of IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY. He is also a member of the International Technology Roadmap for Semiconductors (ITRS) technical working group for Assembly and Packaging (AP).