<table>
<thead>
<tr>
<th>Title</th>
<th>A thermal resilient integration of many-core microprocessors and main memory by 2.5D TSI I/Os</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Wu, Sih-Sian; Wang, Kanwen; Manoj P. D, Sai; Ho, Tsung-Yi; Yu, Mingbin; Yu, Hao</td>
</tr>
<tr>
<td>Date</td>
<td>2014</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10220/25512">http://hdl.handle.net/10220/25512</a></td>
</tr>
</tbody>
</table>
| Rights      | © 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [http://dx.doi.org/10.7873/DAT.2014.190 ].
A Thermal Resilient Integration of Many-core Microprocessors and Main Memory by 2.5D TSI I/Os

Sih-Sian Wu*, Kanwen Wang*, Sai Manoj P. D.*, Tsung-Yi Ho†, Mingbin Yu‡, Hao Yu*

*School of Electrical and Electronic Engineering Nanyang Technological University, Singapore 639798
†National Cheng Kung University, Taiwan
‡Institute of Microelectronics, A*STAR, Singapore

Abstract—One memory-logic-integration design platform is developed in this paper with thermal reliability analysis provided for 2.5D through-silicon-interposer (TSI) and 3D through-silicon-via (TSV) based integrations. Temperature-dependent delay and power models have been developed at microarchitecture level for 2.5D and 3D integrations of many-core microprocessors and main memory, respectively. Experiments are performed by general-purpose benchmarks from SPEC CPU2006 and also cloud-oriented benchmarks from Phoenix with the following observations. The memory-logic integration by 3D RC-interconnected TSV I/Os can result in thermal runaway failures due to strong electrical-thermal couplings. On the other hand, the one by 2.5D transmission-line-interconnected TSI I/Os has shown almost the same energy efficiency and better thermal resilience.

I. INTRODUCTION

Many-core microprocessors with integrated main memory have become the recent interest for the design of high-performance servers. The 2D integration has low bandwidth, long latency and hence poor I/O energy efficiency, since limited data are transferred with long time. On the other hand, the 3D integration [1] by vertical stacking is one promising solution for memory-logic integration by through-silicon-via (TSV), which can significantly reduce the communication latency, improve communication bandwidth and hence result in high I/O energy efficiency.

However, the vertical stacking by TSV results in a long heat dissipation path, which significantly degrades the thermal reliability [2]. From device perspective, due to the isolation material (liner) that surrounds the TSV where the heat is accumulated, it can introduce significant delay [3]. From system perspective, memory-logic integration with high-density DRAMs has significant leakage current, which can be coupled with temperature to form a positive-feedback loop resulting in thermal runaway failure [4].

Instead of stacking memory and logic at different layers with TSV I/Os, the recent introduction of through-silicon-interposer (TSI) [5] provides a 2.5D solution for memory-logic integration. In contrast to TSV which is short RC-interconnect for inter-layer communication, TSI is usually designed as transmission line (T-line) targeted for high-speed long-distance communication between main memory and cores. Compared to the RC-interconnect with repeaters, 2D single-ended T-line (STL) or differential T-line (DTL) [6] with current-mode-logic (CML) buffers [7] have demonstrated better latency, power and bandwidth performance but with large area overhead. However the TSI based T-line can be designed through and under the common substrate, so high performance yet low area overhead memory-logic integration can be achieved. What is more, as memory and logic components are spread on the common substrate that is close to heat-sink, the thermal reliability of 2.5D integration by TSI I/Os can be better than 3D integration by TSV I/Os.

In this paper, we build up a memory-logic-integration design platform to evaluate thermal reliability of the integrated many-core microprocessors and main memory by 2.5D TSI-based I/Os and 3D TSV-based I/Os. With consideration of electrical-thermal coupling, detailed delay and power models of 2.5D and 3D I/Os are introduced into microarchitecture-level cycle-accurate simulators. With the use of general-purpose SPEC CPU2006 [8] and cloud-oriented Phoenix [9] benchmarks, the following results are observed. As for energy efficiency, the 2.5D integration shows almost the same efficiency as 3D integration when the temperature is high. Moreover, thermal runaway failure is prone to be observed in 3D integration by TSV I/Os with more than four layers.
II. 2.5D/3D MICROARCHITECTURE PLATFORM

Figure 1 shows the platform of many-core microprocessors and main memory by 2.5D and 3D integrations. TSV and TSI are two I/O solutions considered in this memory-logic-integration design.

As shown in Figure 1(a), face-to-back bonding is applied between layers. As heat-sink is on the bottom, there is long heat dissipation path for 3D stacking. So the accumulated heat can affect the TSV I/Os performance severely. In contrast, for 2.5D bonding with TSI I/Os, four main memory banks and cores can be spread uniformly on substrate with close distance to heat-sink, as shown in Figure 1(b). As such, the thermal reliability concern would be relaxed in the 2.5D architecture. TSV I/Os in Figure 1(c) are used for vertical connection of chips on different levels of the same IC [10], which are usually designed as RC-interconnects. TSV I/Os can also be used for signal transmission and power propagation. While in Figure 1(d), TSI I/Os are utilized for horizontal connection of two chips on the same level but on the different ICs [5]. The substrate is drilled and TSI I/Os are formed underneath to connect two dies. TSI I/Os for 2.5D bonding are usually designed as transmission lines. In the following, we will provide detailed physical models of 3D TSV and 2.5D TSI I/Os.

III. 2.5D/3D DELAY AND BANDWIDTH MODEL

The first thermal reliability concern of memory-logic by 2.5D or 3D integration is electrical-thermal coupling to delay.

A. Nonlinear Temperature Dependent Capacitor

The traditional 2D interconnect is modeled as RC-circuit with linear dependence to temperature. As shown in Figure 1(c), due to existence of liner material around TSV metal, depletion region is formed [5], [11] when signal voltage is applied across TSV. As such, TSV capacitance $C_T$ and TSV resistance $R_T$ can be modeled by

$$\frac{1}{C_T} = \frac{1}{C_{ao}} + \frac{1}{C_{dep}}; \quad R_T = \frac{\rho d}{\pi a^2}$$

(1)

- $C_{ao}$ and $C_{dep}$ are the height of TSV and the outer radius of TSV metal.
- $\rho$, $d$, $a$ are the height of TSV and the outer radius of TSV metal.

Note that the presence of liner material around TSV forms a nonlinear capacitor against biasing voltage and temperature. The temperature dependent TSV model can be given as

$$R_T = R_0(1 + \alpha(T - T_0)); \quad C_T = C_0 + \beta_1 T + \beta_2 T^2$$

(2)

where $R_0$ and $C_0$ are resistance and capacitance of TSV or TSI at room temperature $T_0$.

$\alpha$ is the temperature dependent coefficient for resistance,

$\beta_1$, $\beta_2$ are temperature dependent first and second order coefficients of capacitance.

All the coefficients can be characterized from measurement [11].

B. Temperature-dependent Delay Model of TSV

A TSV-based 3D I/O channel with buffers at both ends is shown in Figure 2(a). We use the inverter as the buffer for TSV. The corresponding delay model for one TSV I/O channel with capacitance $C_T$ modeled by (2) is shown in Figure 2(b). Based on the TSV resistance $R_T$ and capacitance $C_T$ in (1), the delay of one TSV I/O channel is given by

$$D_{Tsv} = R_0(1 + \alpha T)\beta_2 T^2 + R_0[(1 -\alpha T_0)\beta_1 T + \alpha (D_0 + R_0 C_0) + (1 -\alpha T_0)R_0 \beta_1 T + (1 -\alpha T_0)(R_0 C_0 + D_0)]$$

(3)
IV. 2.5D/3D POWER AND THERMAL MODEL

A. Power Models

As shown in Figure 1, the 2.5D and 3D integrations consist of cores, memory and I/O channel. The power model of each component needs to be studied for thermal analysis.

1) Core and DRAM Power Model: Core power $P_{\text{core-total}}$ is the sum of dynamic power $P_{\text{core-dyn}}$ and leakage power $P_{\text{core-leak}}$,

$$P_{\text{core-total}} = \eta \cdot C_{\text{core}} \cdot V_{DD} \cdot \Delta V \cdot f + V_{DD} \cdot I_{\text{leakage}}$$

with

$$I_{\text{leakage}} = A \cdot \frac{W}{L} \cdot \eta \cdot v_{f} \cdot (1 - e^{-L \cdot f}) \cdot e^{-V_{DD} / \eta}$$

- where $V_{DD}$ is the supply voltage with $\Delta V$ swing,
- $C_{\text{core}}$ is the core capacitance,
- $\eta$ represents the activity factor and $f$ is the clock frequency,
- $L$ and $W$ are the effective device channel length and width,
- $A$, $v_{f}$ are technology-dependent constants, $ns$ is the subthreshold swing coefficient and $v_{f}$ is the thermal voltage.

The dynamic DRAM power with a data array of $n$ identical banks and each bank has B I/O channels can be given as

$$P_{\text{DRAM-dyn}} = n \cdot B \cdot \eta \cdot C_{\text{channel}} \cdot V_{DD}^2 \cdot f$$

- where $C_{\text{channel}}$ is the channel capacitance.

Assuming that one DRAM memory bank with size of $M$ has leakage current similar to (8), the leakage DRAM power can be given as

$$P_{\text{DRAM-leak}} = n \cdot M \cdot V_{DD} \cdot I_{\text{leakage}}$$

It can be observed that the leakage current can vary significantly with temperature in an exponential fashion which may lead to two consequences. Firstly, the leakage power can dominate the total power. Secondly, it may form a positive feedback loop with temperature that can lead to thermal runaway failure.

2) I/O Power Model: The dynamic power of 3D I/O with $N$-channel TSVs and buffers can be given as

$$P_{\text{3D-I/O}} = \eta \cdot N \cdot (C_{T} + C_{L}) \cdot V_{DD}^2 \cdot f$$

- where $\eta$ represents the activity factor,
- $C_{T}$ is the temperature dependent capacitance given in (2),
- $C_{L}$ is the load capacitance of the buffer.

Based on (11), for a TSV I/O channel operating at a high temperature, the dynamic power will go up significantly due to the nonlinear temperature dependence of capacitance. The dynamic power of 2.5D I/O with $N$-channel TSI I/Os and buffers can be calculated as

$$P_{\text{2.5D-I/O}} = \eta \cdot N \cdot V_{DD}^2 \cdot f$$

- where $s$ is the duration of signal pulse, $\eta$ is the activity factor,
- $Z_{0}$ is the characteristic impedance from (5),
- $R_{D}$ is the driver impedance of CML buffer.

By observing (5) and (12), it can be concluded that the power of TSI I/O channel is less dependent on temperature (with square-root dependence) than TSV I/O channel (with quadratic dependence).

B. Thermal Runaway Failure

The second thermal reliability concern is between power and temperature. Under large bandwidth, it is more liable to form a positive feedback loop resulting in thermal runaway failure. The thermal dynamics with heat-sink heat removal ability is [14]

$$C_{TR} \frac{dT}{dt} = P_{\text{thermal}} - \frac{1}{R_{j}} \cdot T - T_{j}$$

- where $P_{\text{thermal}}$ is thermal power,
- $C_{TR}$ is thermal capacitance,
- $R_{j}$ is the thermal resistance path from $g$ chips to the heat-sink.

If the thermal source grows much faster than heat removal ability of heat-sink, temperature will increase exponentially. Thermal runaway temperature $T_{\text{threshold}}$ is temperature at which thermal runaway failure happens. To avoid thermal runaway failure, we can place heatsink closer to processing cores. Since 2.5D integration has a much close heat-removal path to heat-sink, it shows better heat removal ability than 3D integration.

V. EXPERIMENTAL RESULTS

A. Experiment Setup

The system setup for microarchitecture evaluation platform is shown in Table I. The simulation is performed in gem5 [15] with multi-core system set up as 16-core x86 microprocessors. Benchmarks from SPEC CPU2006 [8] and Phoenix [9] are used. The performance results from gem5 are then sent to McPAT [16] to analyze cores scaled at 32nm. The area estimations are obtained from simulation results.

In addition, the number of I/O channels for both 2.5D and 3D integrations is 64. The length of TSV I/Os is 50um, which is the distance of adjacent layers. The length of TSI I/Os is 1.5mm, which is the distance between two ICs. Then the total area is about 288mm$^2$ for 2.5D integration and 64mm$^2$ for 3D integration. Moreover, a thermal simulator [18] is employed to provide temperature profiles for thermal runaway failure observations in both integrations.

B. Device and System Model Results

1) Device Model Results: We first study the delay and power relations with interconnect length in Figure 4. At 25°C, the delay of TSV is around 22.50ps in 50µm and 40.50ps in 90µm with linear relation as well as power with 0.11mW in 50µm and 0.20mW in 90µm. While at 25°C, the delay of TSI is around 34.52ps in 1.5mm and 60.04ps in 3.0mm length, also with linear relation. The power of TSI is 4.11mW in all length. Compared to the TSV I/Os, the power of TSI I/Os shows little dependence on the interconnect length.

Next, the delay and power relations with temperature is shown in Figure 5. The nonlinear temperature-dependent capacitance has more effect on TSV I/Os. For example, the delays of TSV and TSI are 22.50ps and 34.52ps at 25°C. When the temperature rises to 130°C, the delays become 31.95ps and 35.53ps, with increase of 42% and 3%, respectively. As such, TSV I/Os delay is more sensitive to the temperature. With the same temperature transition, the power of TSV and TSI show increase change by 20% and 2%.

2) System Model Results: We further study the system level performance of two integrations with system power breakdown shown in Figure 6. DRAM power is the dominant factor in the system. For example in Figure 6(b), the DRAM power accounts for 65.33% with the power core accounts for just 34.24%. The TSI I/Os power is less than 6% of whole system power consumption and it is insensitive to the temperature. While the TSV I/Os power contributes less than 4% of the whole system power consumption. With the temperature rising, TSV I/Os power percentage will also increase.
The bandwidth in Figure 7. The bandwidth can be adjusted by varying the widths: (a) at 25°C; (b) 2.5D integration at 120°C; (c) 3D integration at 25°C; (d) 3D integration at 120°C.

Fig. 6: Power breakdown of (a) 2.5D integration at 25°C; (b) 2.5D integration at 120°C; (c) 3D integration at 25°C; (d) 3D integration at 120°C.

Fig. 7: Total power and energy efficiency with different bandwidths: (a) at 25°C; (b) at 100°C.

We now present the power and energy efficiency relation with bandwidth in Figure 7. The bandwidth can be adjusted by varying the number of I/O channels. On average, 2.5D integration consumes 62% more power than 3D integration at 25°C. When the temperature rises to 100°C, it consumes 10% more power. If heat dissipation is not well designed, it will form a positive feedback loop between leakage power and temperature resulting thermal runaway failure. Note that the delay of the TSVs will be greatly affected when the temperature is high, thus decrease the bandwidth. Here the energy efficiency is defined as energy consumption per bit. It can be seen that energy efficiency will decrease as the number of channel increases. On average, 2.5D integration consumes 21% more energy per bit than 3D integration at 25°C. But when the temperature rises to 100°C, 2.5D integration achieves almost the same energy efficiency as 3D integration.

In the following, we will explore the thermal runaway issue in both integrations. The initial temperature is set at 25°C. The heat-sink size is 4.0cm × 4.0cm for 2.5D integration and 2.0cm × 2.0cm for 3D integration, both with heat-removal resistance of 4.6K/W. The system temperature trend is shown in Figure 8. When the temperature goes beyond the threshold temperature (100°C), thermal runaway failure happens. The temperature of 2.5D integration is maintained between 50°C and 70°C. For 3D integration, we vary the number of memory layers to see its heat dissipation performance. We can see that it remains stable below 4 layers. With 5 layers, the temperature rising trends can be observed after 6 millions execution cycles, which rises quickly beyond 100°C. Note that the number of memory layers will affect the system performance. As such, high performance 3D system is more liable to the risk of thermal runaway issue. In this case, 4 layers setting is the best from thermal perspective.

The low-area-overhead 2.5D TSI based integration can achieve similar bandwidth and power performance as 3D TSV based integration.

VI. CONCLUSIONS

Thermal resilient memory-logic-integration is studied in this paper by 2.5D TSI I/Os with comparison to 3D TSV I/Os. Detailed electrical-thermal coupled delay and thermal models are developed. The 3D TSV RC-interconnect based I/Os show better delay and bandwidth performance but are more sensitive to the temperature when considering poor thermal-conductive isolation layer and leakage-intensive memory. The 2.5D TSI T-line based I/Os have much less electrical-thermal coupling to delay and power, hence are more resilient to temperature-dependent thermal runaway failure. Moreover, the low-area-overhead 2.5D TSI based integration can achieve similar bandwidth and power performance as 3D TSV based integration.

REFERENCES