Enabling High-density Heterogeneous Integration using Wafer-scale Chiplet Reconstitution Technology

Ashita Victor and Muhannad S. Bakir
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
Email: avictor8@gatech.edu

Abstract

As the demand for high-performance computing continues to grow, the integration of more and more chiplets into a package has become increasingly necessary. In this paper, we present a wafer-scale chiplet reconstitution technology that leverages low-temperature silicon dioxide to encapsulate a dense array of chiplets. To achieve high-density integration and ensure complete oxide encapsulation, a study is carried out to determine the minimum chiplet-to-chiplet gap that is required for a specified thickness of the chiplet. The study shows that a minimum ratio of ~ 1:2 (chiplet height: chiplet gap) is required to fill the gaps between the adjacent chiplets. Based on this observation, a test structure is designed consisting of a dense array of 150 μm x 150 μm passive chiplets that are 12 μm-thick with a chiplet-to-chiplet gap of 40 μm. The designed chiplet tier is fully encapsulated with 18 μm-thick low-temperature silicon dioxide. In addition, a compact circuit model for preliminary analysis of digital die-to-die signaling of the chiplet reconstitution technology is presented. The latency and energy dissipation of this model are evaluated using SPICE simulations.

Key words

Chiplets, Encapsulation, SPICE modeling, Wafer-scale chiplet reconstitution technology.

I. Introduction

There has been a continuous push to develop next-generation packaging solutions to satisfy the growing demand for high-performance computing (HPC). These solutions are expected to offer high bandwidth density, increased energy efficiency, and lower latency [1], [2]. Chiplet-based packaging solutions are key to enabling power-efficient performance [3]. Some examples of commercial chiplet-based solutions include Intel's Ponte Vecchio, AMD's EPYC, Tesla's Dojo, etc. These examples demonstrate that the number of chiplets in a package is increasing, which makes it a challenge to achieve near-monolithic integration [4].

Wafer-level packaging (WLP) techniques can help tackle this challenge by enabling reduced chiplet-to-chiplet interconnect lengths and increased interconnect density [1]. Conventional wafer-level packages often use organic materials such as epoxy mold compound (EMC) to encapsulate the known good dice (KGD) and form a reconstituted wafer. While EMC is a good low-loss material, the mismatch in the coefficient of thermal expansion (CTE) of the EMC and the chiplet results in die-shifts/misalignments and warpage. This limits the integration density that can be achieved [5].

To overcome these limitations, there has been a shift towards alternative materials for encapsulation/embedding of chiplets for wafer-scale packaging. Glass-based substrates have gained significant traction for embedding chiplets [6-7], owing to their superior mechanical and thermal stability, which results in higher interconnect density. Unlike organic substrates, glass has low CTE mismatch with the Si chiplet and is a suitable material for high-frequency applications due to its low-loss properties [7]. In addition, the ultra-low flatness of glass substrates results in an improved depth of focus for lithography [8].

For encapsulation-based wafer scale packaging techniques, inorganic materials like silicon dioxide (SiO_2) have been explored as potential options for encapsulating chiplets. The use of silicon dioxide as an encapsulant offers several advantages. Firstly, it is a low-loss dielectric, which is beneficial for high-speed signaling. Secondly, the possibility of die-shifts or misalignments is reduced as curing is not required. Lastly, it is compatible with standard silicon processes, which allows for post-processing of chiplets using conventional CMOS processes. This blurs the lines between package processing and silicon processing, making it a

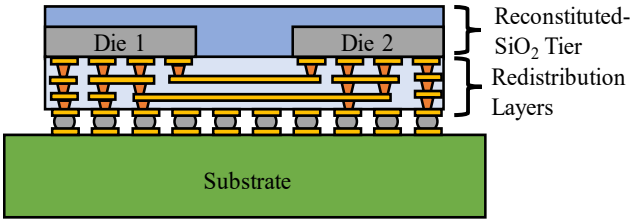


Fig. 1. Schematic of the low-temperature SiO_2-based chiplet reconstitution technology.

highly desirable feature of silicon dioxide as an encapsulant material [4], [9], [10].

Previously, we proposed the encapsulation of chiplets in low-temperature silicon dioxide (Fig.1), known as integrated chiplet encapsulation (ICE), to enable high-density heterogeneous integration (HI) [9]. In [10], we also proposed the idea of integrating a copper heat spreader with the reconstituted-SiO_2 tier to enhance the thermal properties of the package. Similarly, Intel proposed a new high-density HI architecture called quasi-monolithic chips (QMC), incorporating silicon dioxide as an encapsulant material [4].

To achieve near-monolithic integration of the chiplets, it is essential that the chiplets are placed in close proximity to each other. In this paper, we present a testbed to determine the minimum chiplet-to-chiplet gap required to fully encapsulate a dense array of passive chiplets using low-temperature (100 ºC) silicon dioxide. Based on the key findings from this testbed, we demonstrate a single reconstituted-SiO_2 chiplet tier that exhibits complete oxide infill between adjacent chiplets. This paves the path for further post-processing on the tier, ultimately enabling high-density interconnections for efficient die-to-die signaling. Preliminary analysis of die-to-die signaling is carried out by designing compact circuit models and performing SPICE simulations of the wafer-scale chiplet reconstitution technology. These models are evaluated in terms of latency and energy-per-bit.

II. Fabrication Testbeds for SiO_2-based Chiplet Reconstitution Technology

A. Overview of Fabrication Process Flow

For this packaging technology, the fabrication process flow is similar to a chip-first RDL (redistribution layer) last fan-out wafer-level packaging (FOWLP) process. The overview of the fabrication process flow for a single reconstituted SiO_2 chiplet tier is shown in Fig. 2.

The process flow begins by fabricating a dense array of 150 μm x 150 μm passive chiplets with chiplet-to-chiplet gaps determined by the thickness of the chiplets. These passive chiplets are fabricated on Si coupons (20 mm x 20 mm), which are temporarily bonded to a Si carrier wafer using an adhesive. The coupons have 2 μm of PECVD (plasma enhanced chemical vapor deposition) silicon dioxide deposited on the bottom. This is to ensure that during the formation of the chiplets, the adhesive is not directly exposed to plasma. Similarly, the carrier wafer has ~ 4 μm of PECVD oxide to act as a stop layer during the silicon etching processes. After temporarily bonding the coupons, they are thinned down to approximately 10 – 20 μm using inductive coupled plasma reactive ion etching (ICP-RIE) at an etch rate

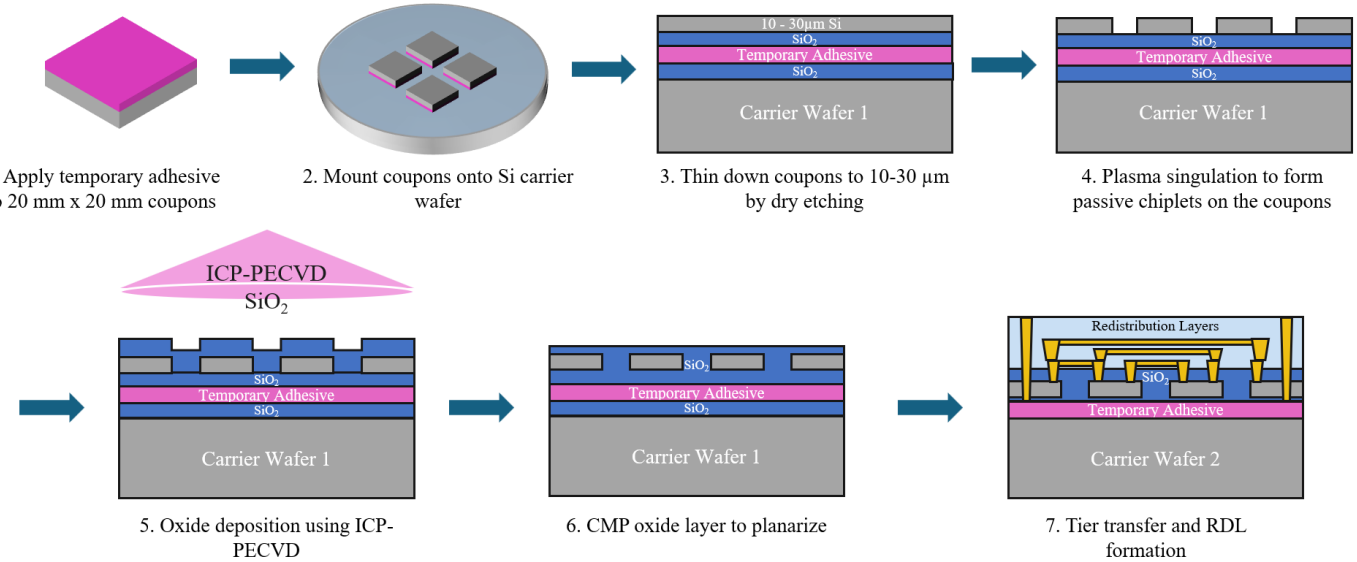


Fig. 2. Overview of the fabrication process flow for a single SiO_2-based reconstituted chiplet tier.

of 6 – 7 μm/min. Next, the chiplets are formed onto the thinned silicon coupon by photolithographic processes followed by a plasma singulation step using a deep reactive ion etching (DRIE) equipment [10].

Finally, the densely spaced chiplets are encapsulated by low-temperature silicon dioxide, which is deposited using an inductive coupled plasma enhanced chemical vapor deposition (ICP-PECVD) equipment. This deposition method helps achieve good quality oxide at lower temperatures (100 ºC). The average deposition rate of this low-temperature oxide is ~ 4.2 μm/hr [9], [10]. The thickness of the deposited oxide is also determined based on the chiplet thickness and the chiplet-to-chiplet gap.

B. Evaluation of Minimum Chiplet-to-Chiplet Gap to Achieve High-density Integration

To achieve high-density chiplet integration, it is crucial to minimize gaps between adjacent chiplets. This results in shorter interconnect lengths, enabling low-latency, high-bandwidth die-to-die signaling [1].

Previously, it has been demonstrated that for a dense array of 150 μm x 150 μm, passive chiplets that are 30 μm-thick with a chiplet-to-chiplet gap of 10 μm, complete oxide encapsulation was not achieved. This is because the oxide is much thicker at the chiplet edges as compared to the sidewalls, also known as oxide bread-loafing. Therefore, when a very thick oxide layer is deposited at a fast deposition rate, it can cause the regions between adjacent chiplets where the chiplet-to-chiplet gap is much smaller than the chiplet thickness, to be pinched off [8]. This issue can be mitigated by increasing the gaps between adjacent chiplets. However, increasing the gaps arbitrarily will increase the length of the interconnects and will result in increased latency and reduced bandwidth.

TABLE I
Summary of Key Observations to Determine Minimum Chiplet-To-Chiplet Gap

Ratio of Chiplet Gap to Chiplet Height	***Thickness of Oxide filled in the Gap (μm)***	***Observation***
0.33:1	< 1 μm	Complete oxide pinch-off, so no oxide infill
0.67:1	< 3 μm	Oxide is close to pinching off, resulting in reduced infill
1:1	< 5 μm	Thickness of oxide infill increases, however it may pinch off for complete infill
1.33:1	< 7.5 μm	Increased oxide infill
1.7:1	< 10 μm	Increased oxide infill

Thus, determining the minimum chiplet-to-chiplet gap for a specific chiplet thickness is necessary to ensure complete encapsulation while achieving near-monolithic integration. To determine the required gap, a testbed is designed consisting of ~ 30 μm deep trenches etched in silicon using DRIE to emulate the chiplet thickness. The gaps (width of the trench) varied in the range of 10 – 50 μm. Next, 15 μm of oxide was deposited using ICP-PECVD. Fig. 3 shows the oxide infill for the different gaps.

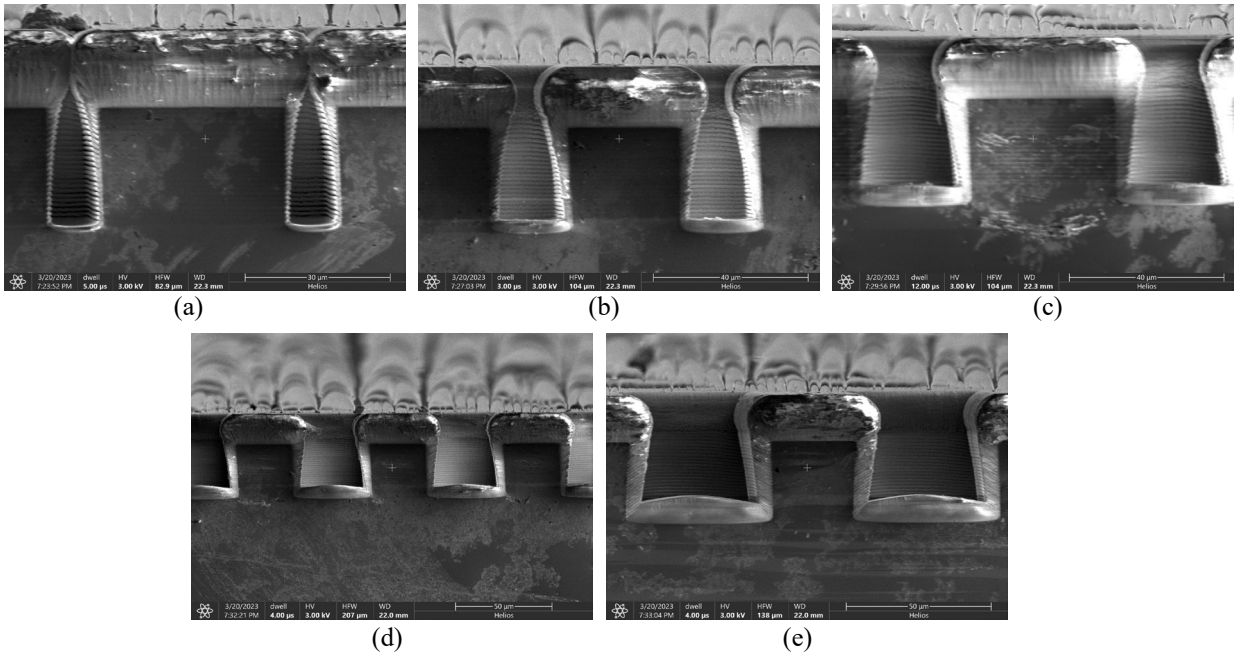


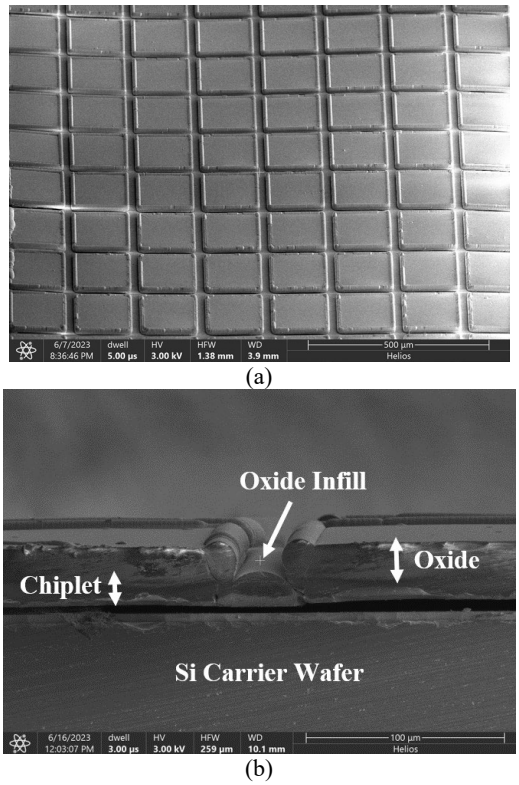
Fig. 3. Evaluation of oxide infill for 30 μm - deep trenches having chiplet-to-chiplet gaps of: (a) 10 μm, (b) 20 μm, (c) 30 μm, (d) 40 μm and (e) 50 μm.

A summary of the observations for each gap is presented in Table 1. The key findings from this testbed showed that the chiplet-to-chiplet gap should be at least twice the chiplet thickness to ensure the gaps between adjacent chiplets are filled with oxide and are not pinched off. An aspect ratio of approximately 1:2 (chiplet thickness: chiplet gap) is observed to fill the gaps with only about 67% of the deposited oxide. The thickness of oxide filled in the gaps increases with wider gaps. However, since the aim is to reduce the spacing between adjacent chiplets to achieve near monolithic integration, it is essential to determine the minimum feasible option.

C. Fabrication of Fully Encapsulated Chiplet Tiers

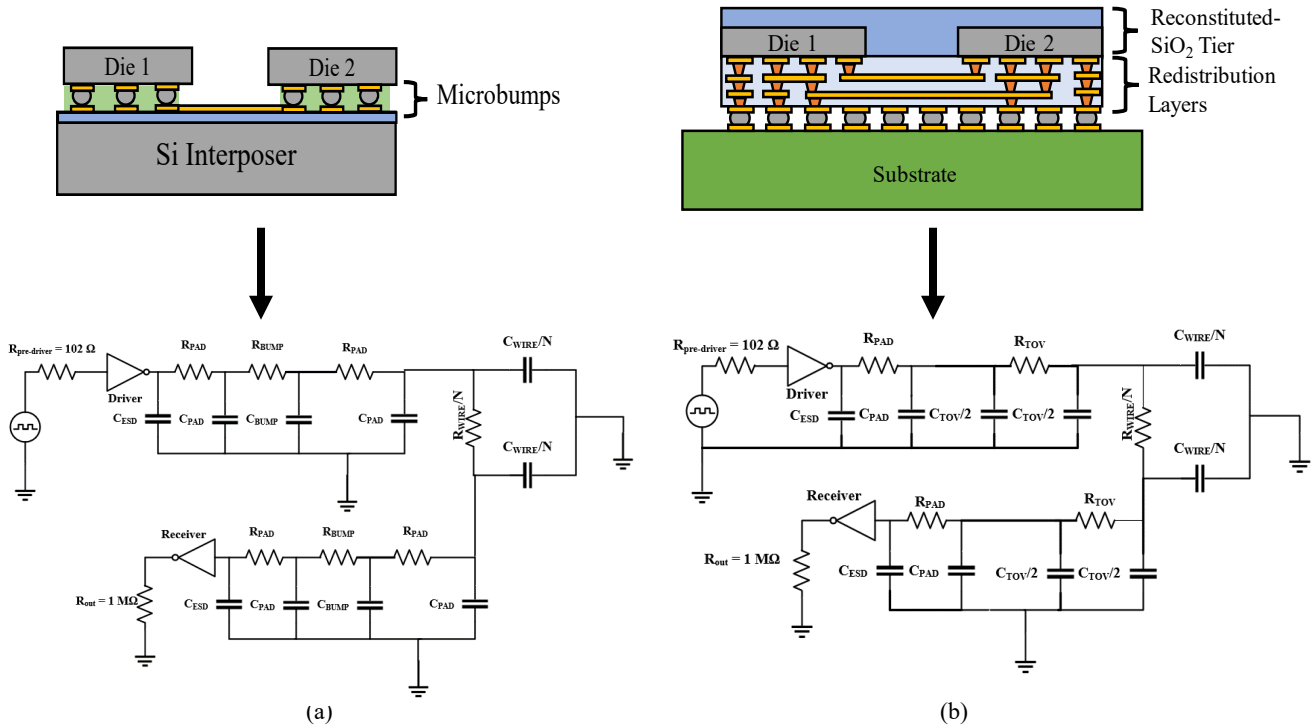
The current testbeds were designed taking into consideration the key observations obtained from the silicon trench-based test structures. The fabrication process flow for these test structures is as described previously. It consists of a dense array of 150 μm x 150 μm passive chiplets that are 12 μm - thick with a chiplet-to-chiplet gap of 40 μm. An 18 μm-thick oxide layer is deposited to encapsulate the dense array of passive chiplets to form a reconstituted-SiO_2 chiplet tier. A cross-sectional view of the chiplet tier with complete oxide infill between chiplets is shown in Fig. 4.

Oxide bread-loafing is observed at the chiplet corners as expected. To carry out further post-processing on these encapsulated chiplet tiers, the deposited oxide must first be planarized by chemical mechanical polishing (CMP),



(a) (b)

Fig. 4. Fully encapsulated reconstituted-SiO_2 chiplet tier. (a) Dense array of chiplets, and (b) Cross-sectional view of complete oxide infill



(a) (b)

Fig. 5. Compact circuit models. (a) Si interposer packaging platform (b) Reconstituted chiplet tier

followed by the transferring of the tier to another carrier wafer to enable the fabrication of high-density redistribution layers.

III. Analysis of Die-to-Die Signaling

In this section, preliminary die-to-die signaling of the wafer-scale chiplet reconstitution technology using a compact circuit model is analyzed. It is evaluated using SPICE simulations in terms of latency and energy-per-bit.

A. Circuit Models and Design Specifications

The compact circuit models for the reconstituted chiplet tier and an example Si interposer-based packaging architecture are shown in Fig. 5. The signal channels consist of I/O drivers and receivers, I/O pads, horizontal chip-to-chip wires, and vertical interconnects; through-oxide-vias (TOV) for the reconstituted chiplet tier and microbumps for the Si-interposer package.

Repeater-based driver and receivers are designed. The number of stages for the inverter chain (repeater) is determined by optimizing the energy-delay product (EDP) for each of the integration schemes, which typically ranges between 1 to 5 stages. The models also include a pre-driver and termination resistor at the receiver side for impedance matching. The pre-driver and termination resistors are assumed to be 102 Ω and 1 MΩ respectively. An electrostatic discharge (ESD) capacitor, 50 fF, is also added to both the driver and receiver ends. For the die-to-die signal wires, a 3-segment pi-model is used to represent the lumped parasitics of the wire [11]. A summary of the different models used for obtaining the circuit elements for each of the components of the signal channel is presented in Table 2.

TABLE II
Circuit Model Elements for Parasitic Extraction

Component	***Equation***
Pad Capacitance (C_{PAD}) [11]	$c_{PAD} = \epsilon_0\epsilon_{ox}\frac{w_{p^2}}{t_{ox}}$ Wp is the pad width and tox is the ILD thickness
Microbump capacitance (C_{bump}) [11]	$C_{bump} = \epsilon_0\epsilon_{UF}\frac{2\pi H_{bump}}{arcosh\left(\frac{P_{bump}}{D_{bump}}\right)}$ H_{bump} is the height, D_{bump} is the diameter and P_{bump} is the pitch of the microbumps
TOV capacitance (C_{TOV}) [12], [13]	$C_{coup} = \epsilon_0\epsilon_{underfill}\frac{2\pi L_{TOV}}{arcosh\left(\frac{S_{TOV}}{D_{TOV}}\right)}$ C_{coup} is the via-via coupling capacitance, L_{TOV} is the length of the via, S_{TOV} is the spacing between vias and D_{TOV} is the diameter of the vias $C_{self} = q_{ss}\epsilon\sqrt{4\pi A_s}$ C_{self} is the self-capacitance for a through-oxide-via, q_{ss} (~1) is constant that depends on the shape of the via and A_s is the surface area of the via
Wire to wire capacitance (C_{Wire}) [11]	$c_{wire} = \epsilon_0\epsilon_{ox}\frac{t_{ox}\times L}{P}$ P and L are the pitch and length of the wire
Wire to substrate capacitance (C_{Wsub}) [11]	$C_{Wsub} = \epsilon_0\epsilon_{ox}\frac{W\times L}{t_{ox}}$ W and L are the width and length of the wire

B. Evaluation of Latency and Energy-per-bit (EPB)

The circuit model of the reconstituted chiplet tier is simulated using Synopsys HSPICE to obtain the 50% - 50% propagation delay and the total energy dissipation of the die-to-die signal channels. The model is simulated with the 45 nm device model from the ASU PTM HP library.

Fig. 6 shows the delay and energy trends using the 45 nm PDK for varying signal channel lengths. From the results it is observed that as the wire length is scaled down from 1 mm to 200 μm, the latency decreases by 9.3% and the energy dissipation decreases by 26.8%. A summary of the energy

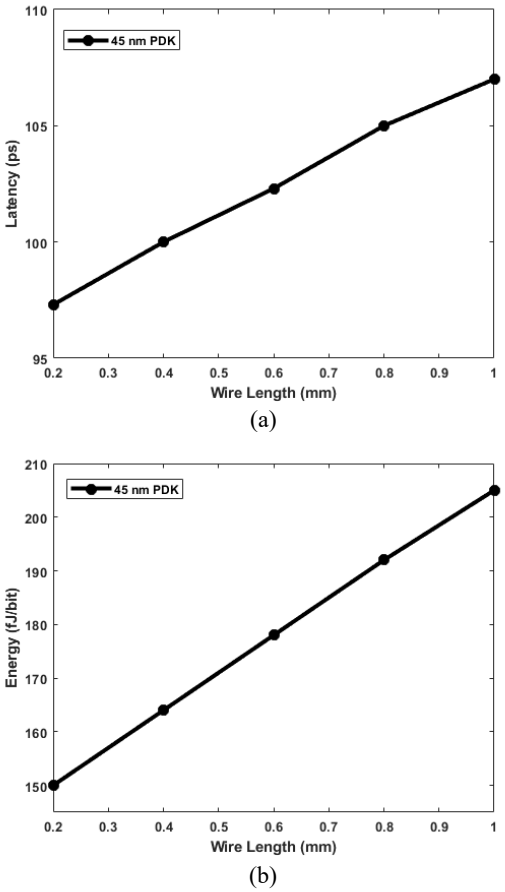


Fig. 6. SPICE simulation for the reconstituted chiplet tier at 45 nm PDK. (a) Latency (b) Energy per bit

TABLE III
Delay and Energy for SiO_2-based Reconstituted Chiplet Tiers

Parameters	*Reconstituted Chiplet Tiers*
Vertical interconnect	Through Dielectric Via
Pitch/Diameter of Vertical Interconnect	2/1 μm
Line/Spacing of Signal Channel	1/1 μm
Dielectric Material	SiO_2
Dielectric Constant	3.9
Delay (@ 45 nm)	107 ps
Energy (@ 45 nm)	205 fJ/bit

and delay results for a wire length of 1 mm is presented in Table 3. To simplify the study, we only analyzed the 45 nm technology node. However, we anticipate that with advanced FinFET nodes, there will be reduced latency due to device performance improvements and increased energy efficiency because of enhanced channel engineering. In addition, the decrease in total capacitance with technology scaling, while I/O capacitance remains constant, highlights the necessity to scale the physical dimensions of the I/O in order to benefit from the technology scaling [11].

IV. Conclusion

This paper discusses a wafer-scale chiplet reconstitution technology that utilizes low-temperature silicon dioxide to encapsulate a dense array of chiplets. Test structures were designed to determine the minimum chiplet-to-chiplet gap that is required for a specified thickness of the chiplet. This study was conducted to ensure that there is complete oxide infill between adjacent chiplets while still achieving high-density chiplet integration. It was concluded that a minimum ratio of ~ 1:2 (chiplet height: chiplet gap) is required to fill the gaps with ~ 67% of the deposited oxide. This implies that for the oxide to be fully flushed with the chiplets, the deposited oxide should be > 1.5x the chiplet thickness.

With these key findings, a fully encapsulated reconstituted chiplet tier is demonstrated. The chiplet tier consisted of a dense array of 150 μm x 150 μm passive chiplets that are 12 μm-thick with a chiplet-to-chiplet gap of 40 μm. Over 18 μm of oxide was deposited to fully fill the gaps between adjacent chiplets.

In addition, a preliminary die-to-die signaling analysis for the reconstituted chiplet tier using a compact circuit model is presented. Using SPICE modeling, the latency and energy dissipation at the 45 nm technology node are determined. As a part of the future work, we aim to benchmark these results with other 2D/2.5D packaging platforms and evaluate the impact of technology scaling on the latency and energy dissipation of the die-to-die channels.

Acknowledgment

This work is supported in part by CHIMES, one of seven centers in JUMP 2.0, a Semiconductor Research Corporation (SRC) program sponsored by DARPA. It is also supported in part by Global Research Collaboration (GRC), an SRC program. This work was performed in part at the Georgia Tech Institute for Matter and Systems, a member of the National Nanotechnology Coordinated Infrastructure (NNCI), which is supported by the National Science Foundation under Grant ECCS-2025462.

References

[1] R. N. Das *et al*., "Extremely Large Area Integrated Circuit (ELAIC): An Advanced Packaging Solution for Chiplets," *2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)*, Orlando, FL, USA, 2023, pp. 258-265, doi: 10.1109/ECTC51909.2023.00051.

[2] R. Agarwal *et al*., "3D Packaging for Heterogeneous Integration," *2022 IEEE 72nd Electronic Components and Technology Conference (ECTC)*, San Diego, CA, USA, 2022, pp. 1103-1107, doi: 10.1109/ECTC51906.2022.00178..

[3] D. D. Sharma *et al.,"* High-performance, power-efficient three-dimensional system-in-package designs with universal chiplet interconnect express", *Nature Electronics*, pp. 1-11, 2024.

[4] A. Elsherbini *et al*., "Enabling Next Generation 3D Heterogeneous Integration Architectures on Intel Process," *2022 International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, 2022, pp. 27.3.1-27.3.4, doi: 10.1109/IEDM45625.2022.10019499.

[5] J. H. Lau *et al*., "Fan-Out Wafer-Level Packaging for Heterogeneous Integration", *IEEE Transactions on Components Packaging and Manufacturing Technology*, vol. 8, no. 9, pp. 1544-1560, Sept. 2018.

[6] S. Ravichandran, V. Smet, M. Swaminathan and R. Tummala, "Demonstration of Glass-based 3D Package Architectures with Embedded Dies for High Performance Computing," *2022 IEEE 72nd Electronic Components and Technology Conference (ECTC)*, San Diego, CA, USA, 2022, pp. 1114-1120, doi: 10.1109/ECTC51906.2022.00180.

[7] X. Jia *et al*., "Antenna-Integrated, Die-Embedded Glass Package for 6G Wireless Applications," *2022 IEEE 72nd Electronic Components and Technology Conference (ECTC)*, San Diego, CA, USA, 2022, pp. 377-383, doi: 10.1109/ECTC51906.2022.00069.

[8] "Intel Unveils Industry-Leading Glass Substrates to Meet Demand for More Powerful Compute," *Intel*. https://www.intel.com/content/www/us/en/newsroom/news/intel-unveils-industry-leading-glass-substrates.html#gs.8g865s (accessed May 01, 2024)

[9] M. J. Li and M. S. Bakir, "3-D Integrated Chiplet Encapsulation (3-D ICE): High-Density Heterogeneous Integration Using SiO2-Reconstituted Tiers", *IEEE Transactions on Components Packaging and Manufacturing Technology*, vol. 11, no. 12, pp. 2242-2245, 2021.

[10] A. Victor, M. Manley, S. Oh and M. S. Bakir, "Reconstituted-SiO2 Tier with Integrated Copper Heat Spreader," *2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)*, Orlando, FL, USA, 2023, pp. 1767-1772, doi: 10.1109/ECTC51909.2023.00302.

[11] Y. Zhang, X. Zhang and M. S. Bakir, "Benchmarking Digital Die-to-Die Channels in 2.5-D and 3-D Heterogeneous Integration Platforms," in *IEEE Transactions on Electron Devices*, vol. 65, no. 12, pp. 5460-5467, Dec. 2018, doi: 10.1109/TED.2018.2876688.

[12] D. K. Cheng et al., *Field and Wave Electromagnetics*, 1989.

[13] C. Xu and K. Banerjee, "Physical Modeling of the Capacitance and Capacitive Coupling Noise of Through-Oxide Vias in FDSOI-Based Ultra-High Density 3-D ICs," in *IEEE Transactions on Electron Devices*, vol. 60, no. 1, pp. 123-131, Jan. 2013, doi: 10.1109/TED.2012.2227966.