Integrating Chiplets using Chips First Ultra-High-Density Fan-out with Maskless Laser Direct Imaging and Adaptive Patterning for High Performance Computing

Benedict San Jose[1], Cliff Sandstrom[1], Erick Talain[1], Jan Kellar[1], Tim Olson[1],
Mary Maye Melgo[2], Rizi Gacho[2], BC Kim[2]
[1]*Deca Technologies Inc., Tempe, AZ, USA*
[2]*nepes hayyim Corporation, Laguna, Philippines*
benedict.san.jose@thinkdeca.com

***Abstract* - Our industry has entered the chiplet era where scaling and performance improvements are enabled through the heterogenous integration of various functions and wafer fab technology nodes through advanced packaging technology.**

Deca's M-Series™ is a chips-first, face-up FOWLP (fan-out wafer level packaging) technology which includes a highly planar surface for RDL build-up where the semiconductor device active surface and vertical sidewalls are fully encapsulated with an epoxy molding compound (EMC) or other dielectric material. Through scaling to 2µm lines & spaces and multiple redistribution layers (RDL), the M-Series provides powerful new possibilities for chiplet-based architectures moving to higher bandwidth interfaces where designers favor multiple parallel interconnects over classic SERDES connections. The desire for thousands of chiplet-to-chiplet connections is driving an unprecedented need for shrinking the device bond pad pitch especially for applications such as high-performance computing.

A major barrier in shrinking the device bond pad pitch is die shift. Die shift is the natural variation in die location within an embedded structure stemming from chip attach, molding, and other process variables. Chips-first fan-out technologies using conventional design methods and mask-based lithography run into barriers in the range of 45µm device bond pad pitch due to the need for large capture pads which account for typical die shift.

The M-Series overcomes die shift using Adaptive Patterning® (AP) with its unique design-during-manufacturing methods and mask-less laser direct imaging (LDI) photolithography. AP enables a precise RDL via connection to each device bond pad or additional layer by precisely aligning a unit-specific pattern to every device. The second-generation M-Series with Adaptive Patterning, or Gen 2, opens up an unprecedented 20µm area array pitch bond pads as a starting point with a roadmap to achieve 10µm bond pad pitch in the near future.

In this paper, we will discuss preliminary Gen 2 design rules and their implementation on a test vehicle jointly produced by Deca and nepes hayyim. The test vehicle has two Chips First chiplet processors and four simulated footprints for high bandwidth memory (HBM) modules which would be mounted Chips Last. In order to enable ultra-high-density interconnect on Gen 2, the design rules allow for a 20µm device bond pad pitch and 2µm lines & spaces. The design and construction of the Gen 2 test vehicle will be examined.

Keywords: Fan-out Wafer Level Packaging; M-Series, Adaptive Pattering; Ultra High-Density Interconnect

I. INTRODUCTION

The semiconductor industry has reached a point where the economic benefits of silicon scaling can no longer be achieved through monolithic integration. As a result, chiplets have become essential in this new era as they enable the integration of various functions and wafer fab technology nodes, leading to cost reduction. Chiplets can now be fabricated using the most suitable wafer fab technology node, ensuring the best performance and favorable commercial terms.

With the industry focus on chiplet integration, and more generally, heterogenous multi-die devices, advanced packaging technologies are becoming the key ingredient for new product designs [1]. This is being driven by the growth of 5G in mobile, artificial intelligence (AI), autonomous driving as well as the internet of things (IOT).

Recently, Heterogenous Integration (HI) modules have been steadily increasing with complexity with the addition of various functional devices and components such as integrated passive devices (IPDs). As an example, we have seen high performance computing (HPC) processors disintegrating the cache memory from the processor in a HI package. This approach has already been proven commercially successful for HPC applications such as the AMD EPYC [2]

In our industry, we are seeing increased levels of integration driving need for tighter pitch. IC transistor density has

2377-5726/23/$31.00 ©2023 IEEE
DOI 10.1109/ECTC51909.2023.00053

increased 520x over the last 20 years, while package density has only increased by 15x (Figure 1). This clearly shows that packaging density is severely lagging behind device density growth [3]. As HI is moving towards integrating more chiplets and functionalities, a high-density interconnect will be needed.

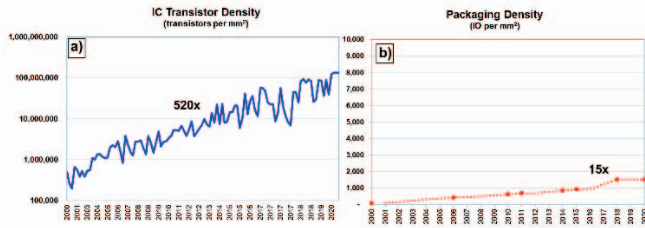

Figure 1: Comparison of IC transistor density increase (a), and packaging density increase (b) from 2000 to 2020

II. The Die Shift problem and its impact on scaling

One of the most significant concerns for embedded die structures is die shift. Embedded die structures have to overcome the impact of variations from multiple sources. During chip attach, the singulated die from one or more source wafers are picked and placed onto a carrier laminated with an adhesive film. Each die has a random X, Y, and rotational offset from its designed position due to mechanical variations and alignment tolerances of the die attach equipment. During compression molding, the flow of the epoxy mold compound, the expansion of the adhesive film, and contraction of the mold compound during curing all contribute to die shift. After molding and post mold curing, the cumulative die-shift from both chip attach and molding. For HI packaging, die shift becomes even more critical as multiple die make the fabrication of die interconnect more difficult to manage. In such a case, one die shifted out of position can lead to loss of the complete package.

A traditional approach to overcome die shift is to add a capture pad to the die contact. In addition, via opening size is reduced to ensure full connection between the via and capture pad. However, this approach limits the minimum achievable bond pad pitch due to the capture pad size, while reduction of the via's electrical contact area may negatively impact electrical performance with higher contact resistance.

Deca has developed AP as a novel approach to solving the die-shift issue in Chips First FOWLP and other embedded die structures. Rather than fight the natural variation inherent to the process, AP generates a unique optimized pattern for each unit in real-time that dynamically accounts for die-shift [4,5,6,7,8].

A demonstration of AP functionality for HI is shown in a simple example in Figure 2. The figure shows an example of a two-die device, one die on the left (DIE 1) and one on the right (DIE 2). The Cu Studs connect through the mold to the die bond pads. The visible RDL pattern was applied without AP, and significant misalignment of the first via layer to the Cu Studs caused by die-shift is visible in the black areas within the circular vias [8].

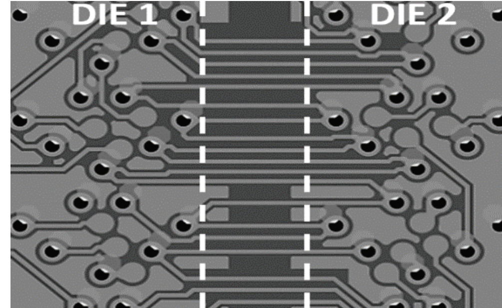

Figure 2. Example of RDL and Via openings not aligned

In Figure 3, the same nominal design is used, however the AP technique of Adaptive Alignment was applied to dynamically align the RDL pattern for DIE 1 and DIE 2 to their respective die locations. DIE 1 and DIE 2 have independent shift variations, so their regions of the RDL pattern are aligned separately. A further AP technique of Adaptive Routing is utilized to dynamically connect the DIE1 and DIE 2 RDL patterns.

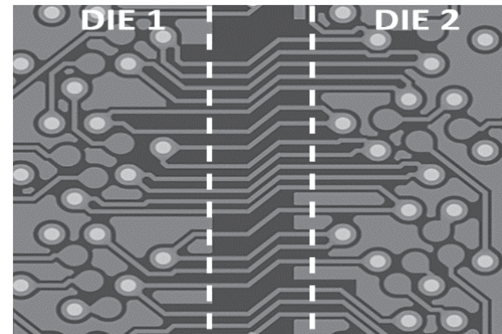

Figure 3. Adaptive Routing between DIE1 and DIE2

III. Scaling to M-Series Gen 2 with Adaptive Patterning

In order to increase density, new interconnect technologies have been developed such as Intel's embedded multi-die interconnect bridge (EMIB), which includes a bridge die within a substrate, scales to a die pad pitch of 45µm with an IO density of 492 IO/mm^2 [9]. Additionally, TSMC's integrated fan-out (InFO-R) supports scaling of die pad pitch down to 40µm achieving an IO density of 545 IO/mm^2 [10].

In comparison, the M-Series planarized structure coupled with AP enables a comparable interface pitch of 45µm with an IO

density of 492 IO/mm^2 in its first generation without the need for complicated bridge chips in substrates such as used by EMIB. Figure 4 compares these pitch capabilities as well as the new M-Series Gen 2 as examples of leading advanced packaging HI solutions.

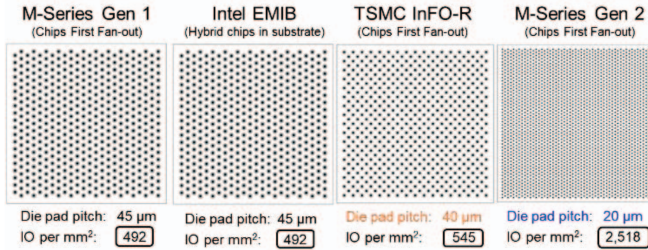

Figure 4: Comparison of competitive interface pitches

M-Series Gen 2 delivers a die pad pitch of 20µm to support an IO density of 2,518 IO/mm^2 utilizing AP techniques and highly planarized build-up surface.

IV. Gen 2 Test Vehicle

A. Design

The Gen 2 test vehicle consists of two processor chiplets and 4 HBM simulated footprints (Figure 5) in an 18×14mm package. The two processor chiplets are interconnected with ultra-high-density routing on the 20µm pitch (Interface A). The HBM footprint die utilizes the JEDEC standard 55µm interface area (Interface B). In contrast, an ultra-high density 20µm pitch on the chiplet processor connects to the standard 55µm pitch on the HBM footprint die (Interface C).

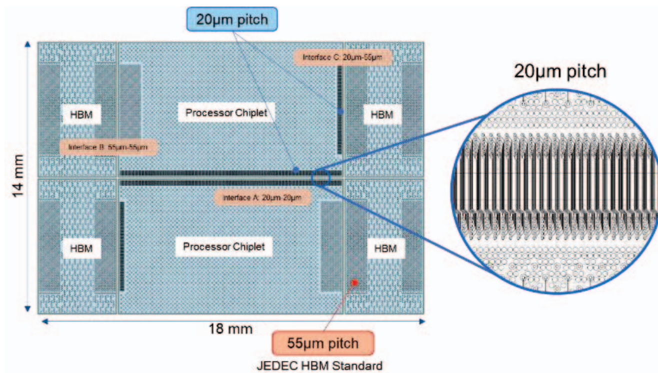

Figure 5. Six die Gen 2 test vehicle

The processor chiplets feature 60µm diameter Cu Studs arranged on a 180µm pitch in the non-interface region. The interface area features a 20µm pitch interface area of the processor chiplets features a 20µm x 30µm pitch array utilizing 12µm diameter Cu Studs (Figure 6a). The HBM footprint die with 55µm pitch is also shown (Figure 6b).

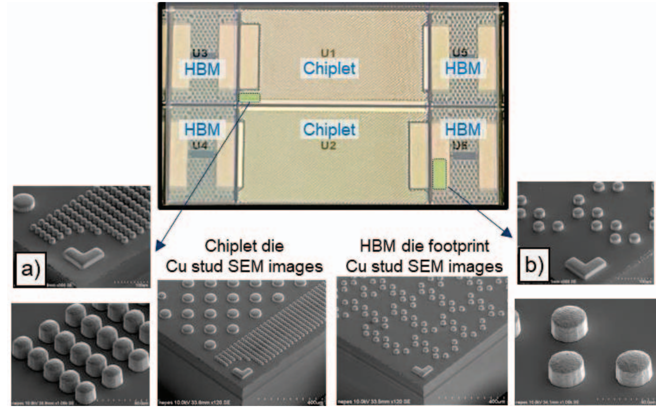

Figure 6. Design and Cu stud images

B. Construction

The test vehicle was produced with the M-Series chips-first, face-up FOWLP process where the semiconductor chips are placed face up on a temporary carrier. During the panelization process, the active surface and vertical sidewalls are fully encapsulated within the dielectric material. A high-level process flow is outlined in Figure 7. The basic flow can be broken down into four key modules: Wafer Prep, Panelization, RDL buildup, and Package finishing.

The wafer prep process entails a standard seed deposition followed by a photo process using a thick dry film or liquid resist to pattern the Cu Studs. Electroplating is utilized to form the Cu Studs. After plating, the resist is stripped and the seed layer etched. Backgrind is then performed to thin the wafer to its target thickness followed by dicing of the wafer to create singulated units.

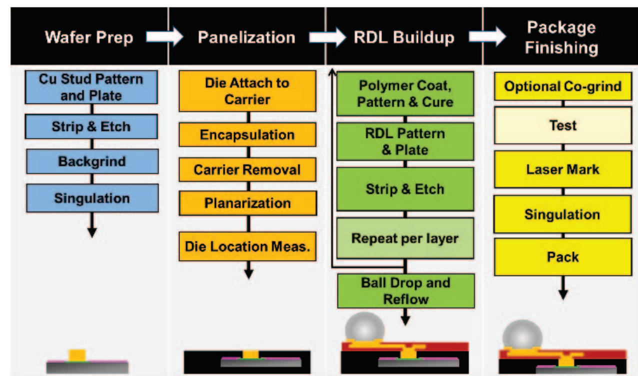

Figure 7. M-Series high-level process flow

During panelization, the chip attach process picks up the die from the singulated film frame and places the die face up on a temporary carrier. After die attach, an encapsulation process is utilized to create a reconstituted 300mm round wafer or 600mm square panel. After encapsulation, the

temporary carrier is removed through a proprietary debond process. The next step is front planarization which reveals the Cu Studs for later connection to the RDL. Following front planarization is the die location measurement process where certain Cu Studs of each individual die are measured to provide input to the AP real time design software which creates a unique optimized GDSII file for each wafer or panel.

The RDL build up process utilizes conventional wafer-level fab processes which include dielectric layers and metal RDL interconnect layers. The number of RDL layers varies by application but is typically one or two for simple structures and three to five for advanced HI packages. The first step of the buildup process includes application of a polyimide (PI) or other suitable dielectric material. LDI is utilized to form via openings in the dielectric layer according to the optimized GDSII file. After dielectric cure, a seed layer is applied to the panel following the same process as the outlined in the Cu stud process. RDL or underbump metal (UBM) layers are formed as per the final desired device structure. Following the RDL and UBM buildup, a standard ball drop, reflow, and flux clean is performed prior to moving to the package finishing process.

An optional co-grind process thins the encapsulant and silicon. The units are then typically prepared for a wafer or panel-based final electrical test. After final test, the devices are marked, typically singulated, and packaged in tape and reel or trays for shipment.

Next generation LDI exposure tools and photoresist that support the Fan-out Gen 2 process have been developed to support 2µm lines & spaces. Polyimide dielectric has also been developed to support 6µm via openings. With LDI, 405nm wavelength light is ideal for system reliability considerations. Therefore, the photoresist and polyimide materials were optimized for exposure using 405nm.

Figure 8a shows 2µm lines and spaces in photoresist exposed with 405nm LDI, while Figure 8b shows a 6µm via formed within the PI dielectric.

In addition to RDL buildup layers, a new Cu Stud photoresist tailored to 405nm was developed to support the Gen2 Cu Stud design rules of 12µm diameter with 8µm spacing.

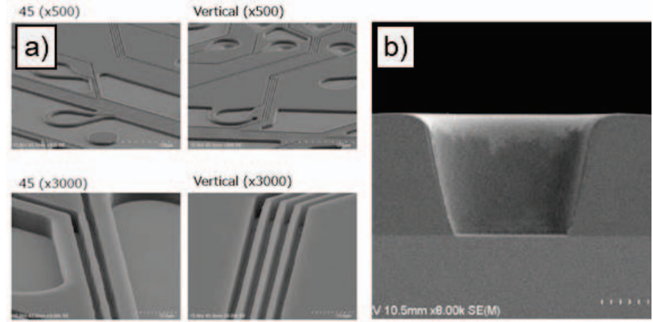

Figure 8. (a) 2µm line & space features (b) 6µm via opening

Figure 9a shows the top view of the processor chiplets and the HBM footprint die after die attach to a 300mm round carrier. Figure 9b shows a close-up isometric view of 9a.

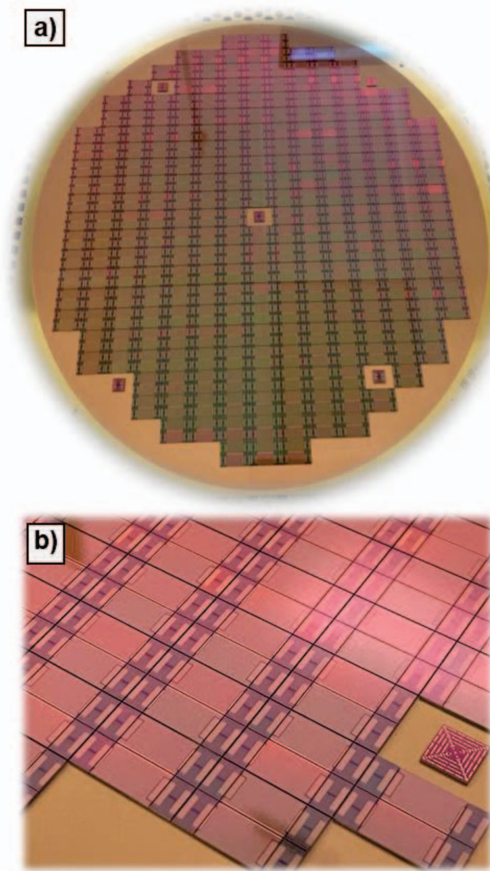

Figure 9. Devices after die attach

The encapsulation process is performed to create a reconstituted wafer. Following the removal of the temporary carrier, the molded wafer is planarized.

During the planarization process, the Cu Studs on the embedded die are revealed. Figure 10a shows the top view of the molded wafer after front planarization. Figure 10b shows a close-up isometric view of 10a.

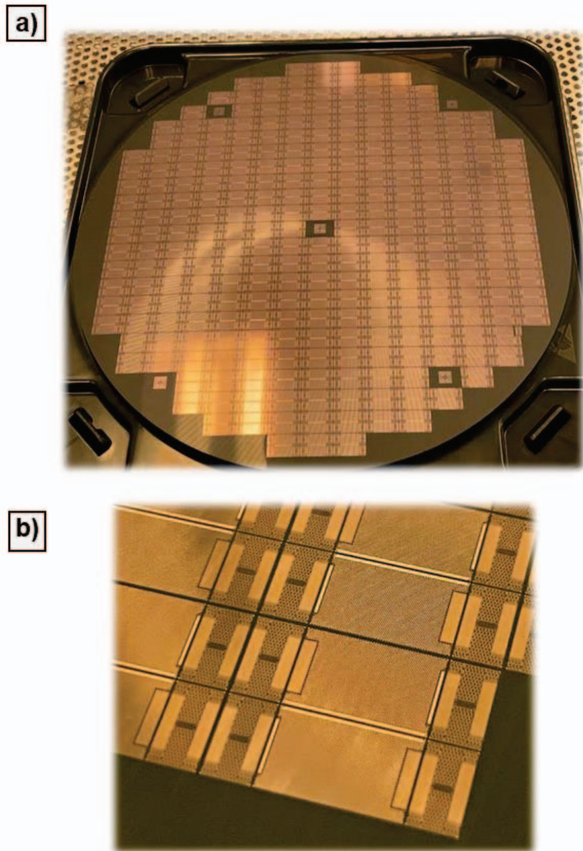

Figure 10. Devices after front planarization

Upon closer examination of the planarized wafer, the different interface densities become evident. Figure 11 shows a schematic of the Gen 2 test vehicle with close-up images highlighting the different regions of the design.

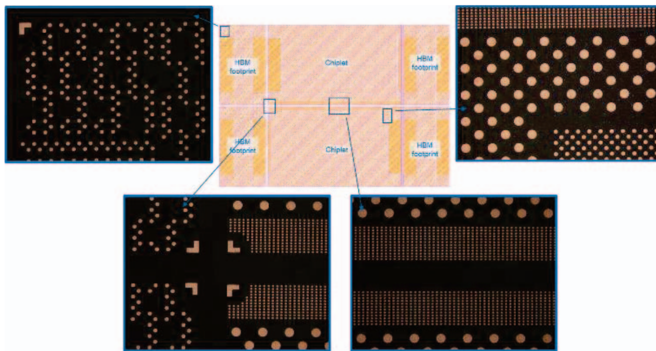

Figure 11. Close-up images of the various interface densities

After planarization, the molded wafers are processed through a die measurement system (DMS). The DMS process utilizes a high-speed optical inspection system with high accuracy. Typical measurement accuracy for each actual die location is better than 500nm at 3σ. Figure 12 shows the molded wafer processing through DMS.

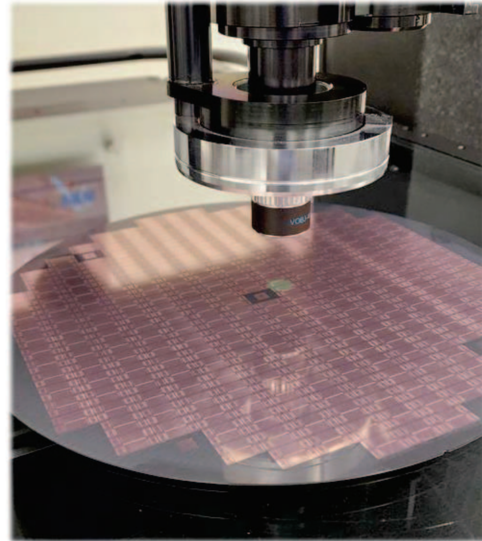

Figure 12. Die measurement process

After DMS, a measurement file for each wafer is sent to the AP system where a unique optimized GDSII file is generated. Figure 13 shows the layout of an actual optimized GDSII file generated by AP.

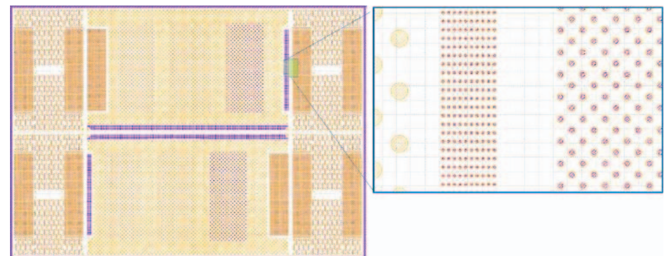

Figure 13. Layout of an optimized GDSII file

Typically, a dielectric is the first layer applied to the reconstituted wafer during the build-up process. Figure 14 shows a polyimide layer where the GDSII via pattern has been exposed opening an electrical connection to the Cu Studs.

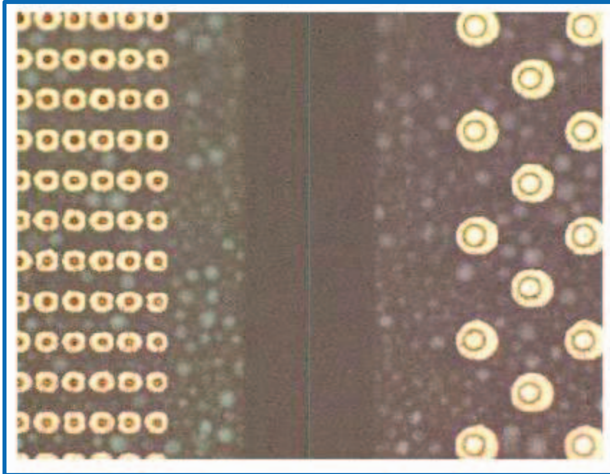

Figure 14. PI via openings over Cu Studs

In subsequent work, RDL build up layers will be added to the structure. We anticipate presenting further results in a future conference.

V. CONCLUSION

In this paper, preliminary M-Series and AP Gen 2 design rules and their implementation on a test vehicle have been presented. The test vehicle includes two Chips First chiplet processors and four simulated footprints for chips last mounting of HBM modules. A 20µm device bond pad pitch with 2µm lines & spaces have been demonstrated using AP for ultra-high density chiplet integration as well as the processor side of an HBM interface.

Acknowledgements

The authors would like to express their appreciation for the support of nepes hayyim R&D in the building of the test vehicle.

References

[1] Yole report: "High-end packaging: Intel and TSMC are competing. What will be the strategy of Samsung and the others?" http://www.yole.fr/iso_upload/News/2020/PR_HIGH_PERFORMANCE_COMPUTING_MarketEvolution_YOLE_SYSTEM_PLUS_CONSULTING_Nov2020.pdf

[2] S. Naffziger, et al., "Pioneering Chiplet Technology and Design for the AMD EPYC™ and Ryzen™ Processor Families: Industrial Product," 2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA), 2021, pp. 57-70, doi: 10.1109/ISCA52012.2021.00014.

[3] (March-April 2020). Post–Moore's Law electronics: now until quantum electronics. Chip Scale Review, Volume 24, Number 2, pp. 5-6.

[4] T. Olson and C. Scanlan, "M-Series Fan-Out with Adaptive Patterning," in *Advances in Embedded and Fan-Out Wafer Level Packaging Technologies*, B. Keser and S. Krohnert, Eds., New Jersey: Wiley, 2018, pp. 117-140.

[5] B. San Jose, C. Sandstrom, J. Kellar, C. Bishop and T. Olson, "Adaptive Patterning and M-Series for High Density Integration.," in *Embedded and Fan-Out Wafer and Panel Level Packaging Technologies for Advanced Application Spaces: High Performance Compute and System-in-Package*, B. Keser and S. Krohnert, Eds., New Jersey: Wiley, 2021, pp. 125-154.

[6] B. Craig and et. al., "Adaptive Patterning Design Methodologies," in *IEEE 66th Electronic Components and Technology Conference (ECTC)*, 2016.

[7] C. Scanlan and e. al., "Adaptive Patterning for Panelized Packaging," in *International Wafer Level Packaging Conference (IWLPC)*, 2012.

[8] E. Hudson and e. al., "Deca & Cadence Breakthrough Heterogeneous Integration Barriers with Adaptive Patterning," in *IEEE 71st Electronic Components and Technology Conference (ECTC)*, 2021.

[9] R. Mahajan et al., "Embedded Multi-die Interconnect Bridge (EMIB): A High Density, High Bandwidth Packaging Interconnect," 2016 IEEE 66th Electronic Components and Technology Conference (ECTC), 2016, pp. 557-565, doi: 10.1109/ECTC.2016.201.

[10] 40um pitch reference; Tom Dillinger, SemiWiki.com, Highlights of the TSMC Symposium, 9/7/2020