Assembly Challenges and Approaches for 2.5D Chiplet Based System

Sharon PeiSiang Lim
*Institute of Microelectronics, Agency for Science, Technology and Research (A*STAR)*
Singapore,Singapore
limps@ime.a-star.edu.sg

Mihai D. Rotaru
*Institute of Microelectronics, Agency for Science, Technology and Research (A*STAR)*
Singapore,Singapore
mihaidr@ime.a-star.edu.sg

Wenwei Seit
*Institute of Microelectronics, Agency for Science, Technology and Research, (A*STAR)*
Singapore,Singapore
seitww@ime.a-star.edu.sg

Hsiao Hsiang Yao
*Institute of Microelectronics, Agency for Science, Technology and Research, (A*STAR)*
Singapore, Singapore
hsiaohy@ime.a-star.edu.sg

2023 IEEE 73rd Electronic Components and Technology Conference (ECTC) | 979-8-3503-3498-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ECTC51909.2023.00031

***Abstract*— In recent years, the semiconductor industry has seen a slow down in Moore's law and witnessed longer time spans between process technology nodes. In addition, the integration of transistors on a monolithic chip is becoming more challenging due to available options for device miniaturization [1,2]. The other issues in chip design and manufacturing are its cost and performance. The application of chip level heterogeneous integration or called the Chiplet technology is a promising solution for this node size limitation. Chiplets are small IC dies with specialized functionality, designed to be combined to make up a larger integrated circuit, following the semiconductor industry's trend of heterogeneous integration [3].**

In this paper, 3 different HDFOWLP MCP is designed integrated with chip-to chip interconnections on the RDL-first FOWLP platform. One of the MCP package is using the Intel functional FinFet 22nm chips.

Keywords— Chiplet, flip chip, RDL-first FOWLP, SIP

I. Introduction

Digital transformation is driving industries to move to 5G, mmWave and automotive technologies. For many years, Moore's law has driven the electronics industry towards the scaling down of process node [4]. However this is no longer applicable due to the limitations of the device and its cost. In order to meet the requirements of advanced devices, chip level and substrate level scaling have been used. An alternative approach is the application of smaller chiplets in the backend packaging solution such as integration of chiplets in SIP format to achieve lower cost, higher performance and volume production [5]. Chiplets are small IC dies with specialized functionality. These are designed to be combined to make up a larger integrated circuit, following the semiconductor industry's trend of heterogeneous integration. Heterogeneous integration allows for integration of different chiplets with separate designs and different manufacturing process nodes to integrate within a single package [6]. Some of the benefits are:

1) Reduced IC design and production costs
2) Provide industry for more advancement in terms of intelligence and more connectivity
3) Higher bandwidth
4) Higher performance at a lower cost
5) Lower power consumption
6) Greater design flexibility

In this paper, 3 different HDFOWLP MCP is designed integrated with chip-to chip interconnections on the RDL-first FOWLP platform. One of the MCP package is using the Intel functional FinFet 22nm chips [7,8]. The dimensions of the 3 different MCP packages are as below:

a) 4 die package size of 13.5x13.5x0.9mm using Intel functional 22nm chips (4x4 mmx0.3mm)
b) 8 die package size of 18x18x0.9mm using daisy chain test chips (5.84x4.14x0.775mm)
c) 3 die package size of 10x6.7x0.9mm using 1.8x3.6x0.775mm test chips

The flip chip dies are attached to under bump metallization (UBM) on the bottom RDL layers and followed by underfill dispensing process. Molding process is done on the RDL-first to encapsulate the flip chip dies. Laser debonding are done to remove the glass carrier and following by cleaning. The molded wafer is singulated into different MCP packages and SnAgCu solderballs are then attached onto bottom UBM of the package.

There are many challenges in the flip chip assembly HDFOWLP MCP. Some of the challenges are: 1) incoming warpage of the RDL-first wafer 2) attachment of the flip chip dies onto bottom RDL layers, 3) poor or open solderjoint interconnections 4) capillary underfill voids. Poor wetting solder interconnection or open solder joints may result if the wafer warpage is too high. In addition, void-free underfill process requires co-optimization of underfill material properties, the dispense process, and cure conditions. Different chuck heating temperature have been evaluated to achieve void-free underfilling process. Wafer level molding is done on the RDL-first wafer followed by solderball attachment process on the 3 singulated HDFOWLP MCP package.

II. Test Vehicle Description and Process Flow

This work demonstrates 3 different HDFOWLP MCP integrated with chip-to chip interconnections on the RDL-first FOWLP platform using flip chip technology. Table 1 shows the dimensions of the different test dies in the 3 MCP packages. The different test dies used in the MCP packages are shown in

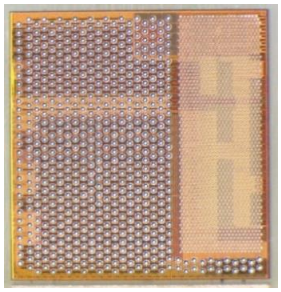
2377-5726/23/$31.00 ©2023 IEEE
DOI 10.1109/ECTC51909.2023.00031

Fig 1. The Intel functional 22nm chips and 8 die chiplet test chips has solder bumps at 55µm and 130µm pitch while the 3 die chiplet test chips has solder bumps at 90µm pitch. The solder alloy for all test chips is SnAg solder.

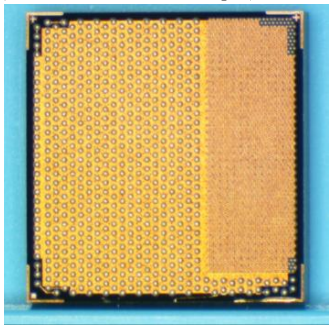
The overall assembly process flow is highlighted in Fig 2. The test dies used in this work are comprises of dies fabricated on a MPW wafer and Intel functional 22nm chips. The RDL and UBM layers are processed on a carrier coated with sacrificial layer for a lower wafer warpage during metallization and passivation layer build-up process. Fig 3. shows the top view of the chiplet packaging design after wafer fabrication. Mechanical dicing is applied for dicing the MPW wafer. The flip chip dies are then attached to under bump metallization (UBM) on the bottom RDL layers and subjected to mass reflow process. This is followed by dispensing of the underfill material to flow under the dies and protect the solderjoints. Molding process is done on the RDL-first wafer to encapsulate the flip chip dies [6]. Laser debonding process is done to release sacrificial layer and remove the glass carrier and following by cleaning. The molded wafer is then singulated into different MCP packages and SnAgCu solderballs are then attached onto bottom UBM of the package.

Table 1: Details of MCP packages

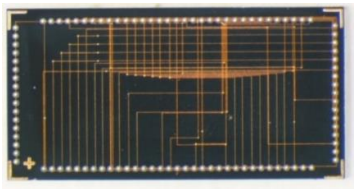
Chip Type	Netflex chiplet (Intel chips)	(Internal daisy chain chips for 8 chiplet MCP)	(Internal daisy chain chips for 3 chiplet MCP)
Die size	4.084x3.924x0.775mm	5.84x4.14x0.775mm	1.8x3.6x0.775mm
I/Os	988	2230	100
Bump pitch	130um for C4 bumps	130um for C4 bumps	90um
	55um for microbumps	55um for microbumps	
Bump structure	Full solder for 130um pitch C4 bumps	Full solder for 130um pitch C4 bumps	Full solder for 90um pitch C4 bumps
	Cu pillar + solder for 55um pitch microbumps	Full solder for 55um pitch microbumps	



(a) Intel functional 22nm chips (4x4 mmx0.3mm)



(b) Daisy chain test chips for 8 chips chiplet (5.84x4.14x0.775mm)



(c) Daisy chain test chips for 3 die chiplet (1.8x3.6x0.775mm)

Fig. 1: Optical image on test dies used in the MCP packages

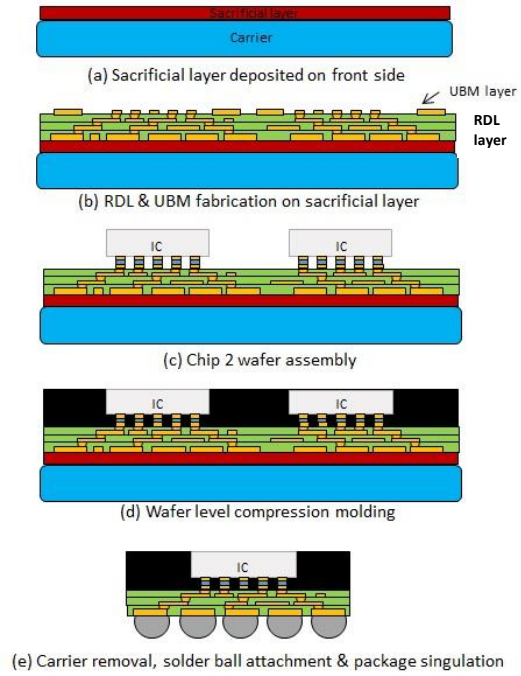


Fig. 2: Overall Assembly Process Flow

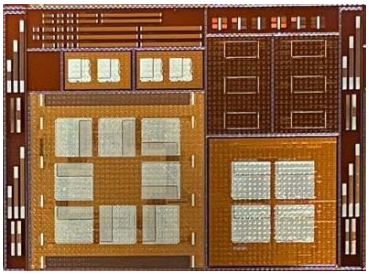
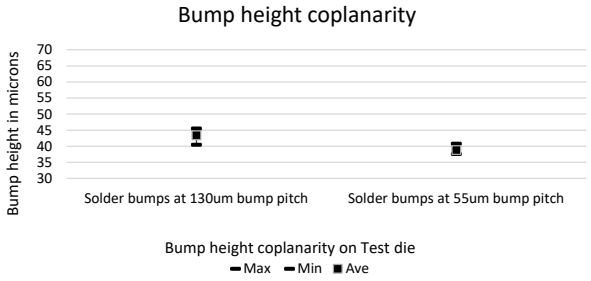


Fig. 3: Top view of the chiplets packaging design layout

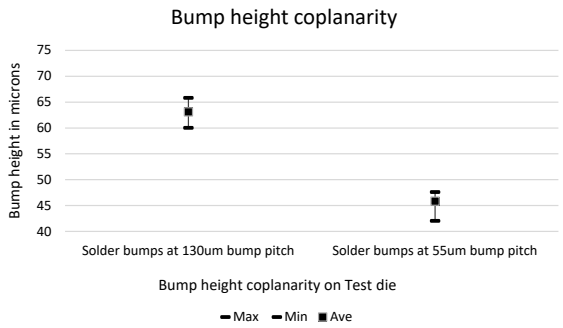
III. Test Die Characterization

Bump characterization such as bump height coplanarity and bump shear test are done on the 3 types of test dies prior to the flip chip assembly process. It is critical to have uniform bump coplanarity as non-uniform bumps may result in weak solder joints and open solder interconnections. Fig 4. Shows the results of the bump coplanarity within the 10% acceptable range. Next solder bump shear is done to inspect the shear strength. This is to prevent any delam on at the bump interface when the samples are subjected to performance testing. The average bump shear strength for the Intel functional chips is 15.22g/mils^2, 4.69g/mils^2 for daisy chain test chips for 8 chips

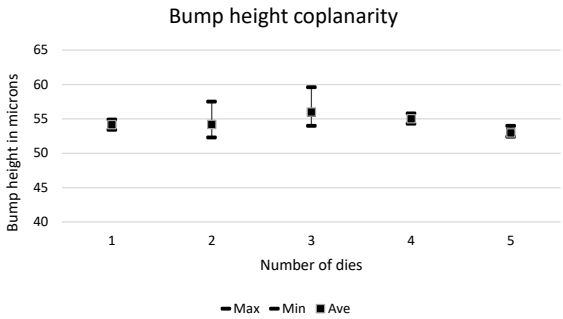
chiplet and 5.78g/mils^2 for the daisy chain test chips for 3 die chiplet. The results of the bump shear is greater than 2.6g/mils^2 of the required specs and the failure mode after is at the solder interface. (Fig 5).



(a) Intel functional 22nm chips

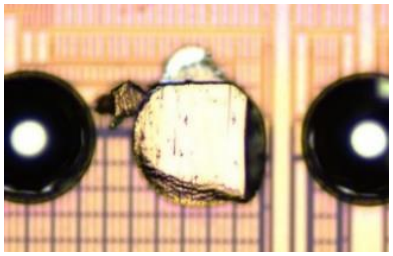


(b) Daisy chain test chips for 8 chips chiplet

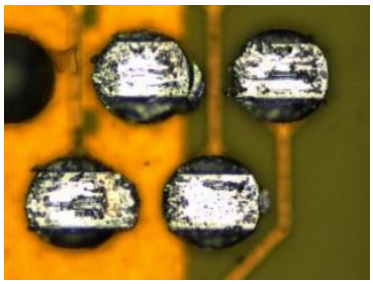


(c) Daisy chain test chips for 3 die chiplet

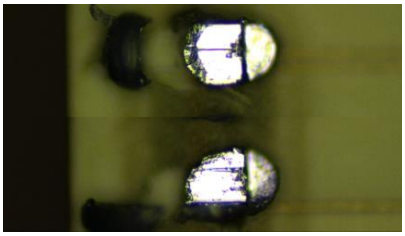
Fig. 4: Bump height coplanarity



(a) Intel functional 22nm chips



(b) Daisy chain test chips for 8 chips chiplet



(c) Daisy chain test chips for 3 die chiplet

Fig. 5 : Bump shear failure mode after shear test

IV. ASSEMBLY PROCESS

A. Chip to wafer bonding

The attachment of the 4 Intel chiplets and the 3 daisy chain chiplets are attached to the bottom RDL wafer and reflow process using a no-clean flux NC26A from Indium. The post reflow flux residue left around the solderjoints does not required cleaning as they are minimal and non-corrosive. The flux dipping thickness is at 10um. The sample then subjected to a conventional reflow oven at a lead-free reflow profile. The peak temperature is ~ 260°C and at a dwell time of 58 secs above the melting point of SnAg solder at 221°C. Fig 6. shows the assembled wafer. The multi-chiplets test vehicle assembled on the RDL 1st FOWLP is shown in Fig 7. On the other hand, the 8 daisy chain chiplets are done using thermocompression process (TCB) due to the fine pitch solder interconnection at 55um and 130um designed respectively on the same chip (Fig. 8). In TCB, a constant force of 15N is applied by the bond head with the test die being rapidly heated up to the liquidus temperature. The die is moved further down in the Z axis to ensure all solder joints are in contact with the UBM pads during the solder melting phase. The bond head then held the package by vacuum to allow the solder to reflow completely and have good wetting. During the solder molten phase, the bond head then moves upwards in Z direction to control the height of solder joint. The die is then cooled rapidly below the solidus temperature. The bond head vacuum is then turned off to release the package. The bond head then moves away as solder joints have solidified. Fig. 9 shows the thermocompression bonding process assembly process flow. A typical TCB bonding profile is shown in Fig. 10.

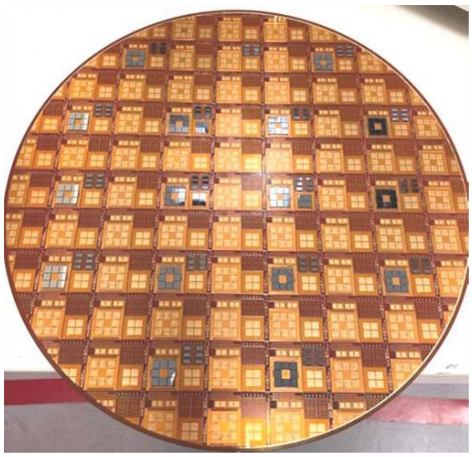


Fig. 6: Assembled FOWLP wafer

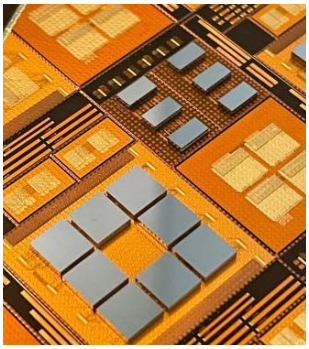


Fig. 7: Multi-chiplets test vehicle on RDL 1st FOWLP

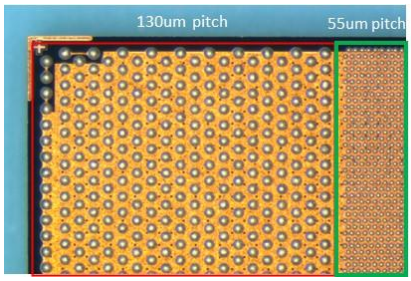


Fig. 8: Daisy chain test chips for 8 chips chiplet

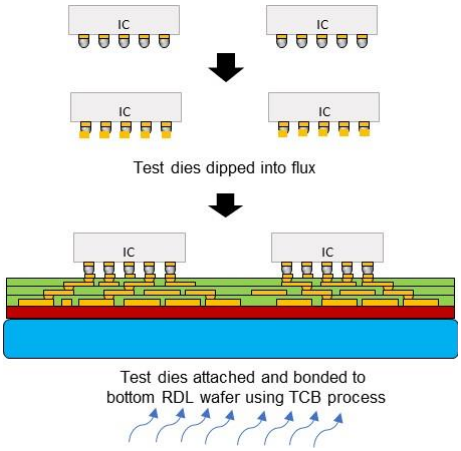


Fig. 9: Assembly process flow for thermocompression bonding process

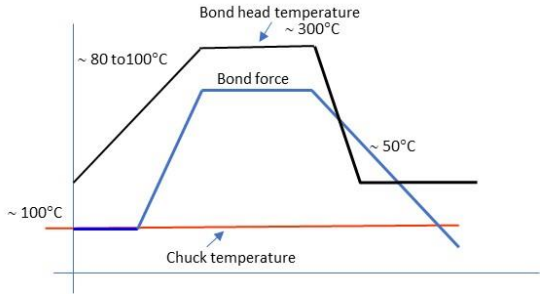


Fig. 10: Schematic drawing of a typical TCB bonding profile

Proper selection of bond force is critical to ensure solder bumps are in contact with bottom RDL UBM pads during the thermocompression process. In the case of too high bonding force, solder bridging will occur between the 130um pitch solder bumps (Fig. 11). Fig. 12 shows the good bonding alignment of the solder bumps to the bump pads without solder bridging.

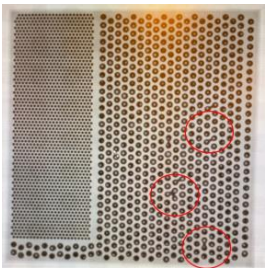


Fig. 11: Xray image on solder bridging after TCB bonding process

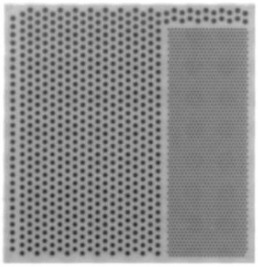


Fig. 12: Xray image on good solder alignment after TCB bonding process

B. Electrical characterization and measurements of multilayer polymer based RDL

The electrical continuity test is conducted on the 8 chiplet package after TCB bonding process to determine the daisy chain results and to ensure good integrity of the solder interconnection. The FOWLP package has a five-layer structure with a metal thickness of 2um and smallest metal features of 2um width and 2um spacing. The stack of the package for the functional 22nm chiplets uses the top metal (m5) as UBM layer where the chiplets are connected through a flip chip interconnection. Layer m4 and m3 are used to distribute the AIB bus that connects the chiplets [7,8]. The parallel bus lines use the minimum features of 2um width and 2um spacing lines to connect the AIB IOs that have a 55um pitch. Layer m2 is used as ground plane, while layer m1 is used to distribute the seven power domains for each chiplet [7,8]. The chiplets have

a dual pitch configuration with 8 channels of the AIB interface IOs located on one side of the die which uses a 55um pitch. The number of the 55um pitch pads is 1609. The rest of the IOs have a 130um pitch. The total number of pads per chiplet is 2230, which are distributed across a 4mmx4mm die area. A drawing illustrating the pads arrangement of the chiplet is given in Fig. 13.

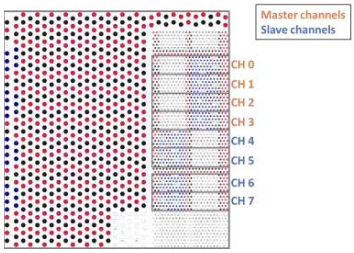


Fig. 13: Pad arrangement of the chiplet

To understand and assess the integrity of the interconnection between the chiplets and the package, a daisy chain dummy chiplets have been designed and fabricated using the exact same size and pads arrangement as shown in Figure X. The dummy chiplets have been assembled in the DC resistance of the daisy chain was measured. Each dummy chiplet had 6 daisy chains, 3 routed through the 130um pitch pads area and 3 routed through the 55um pads. With this approach the integrity of the fine pitch and larger pitch can be assessed independently. Further the dummy chiplets were connected in a network of 8 chiplets to allow assessment of the chiplet to chiplet integrity interconnection (Fig. 14).

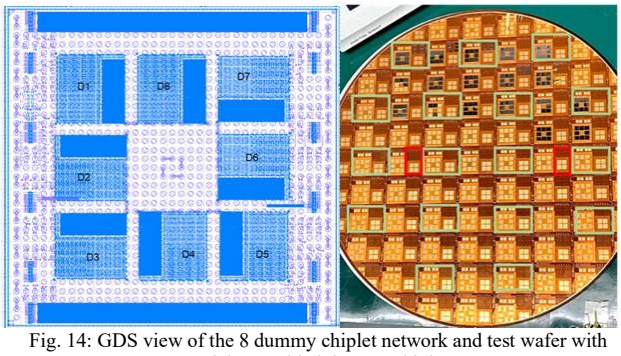


Fig. 14: GDS view of the 8 dummy chiplet network and test wafer with partial assembled dummy chiplets

A test wafer was used to experiment with the assembly conditions and tune assembly process parameters. The dummy daisy chain chiplet were assembled in on three reticles and a DC test was measurement was done on the assembled structures. The data was then feedback into the assembly process to tune and correct the process and its parameters. Then three new reticles were assembled using the new condition and subsequently DC resistance measurements were done again to the last assembled samples. This process was repeated several times to achieve good DC measurement data for all the 6 daisy chains and the daisy chains that connected all the 8 chiplets. An example of the DC chain resistance measurement results after the assembly parameters were optimized are given in Table 2 below. It can be noted that the calculated DC resistance of the daisy chains fit relatively well with the measured data. These results confirm the integrity of the flip chip dual pitch interconnect and allowed us to move to the assembly of the function 4 chiplet structure.

Table 2: Measured DC resistance of the daisy chain for the optimized assembly process

Daisy Chain (pitch of the pads)	Rez (ohms) D1	Rez (ohms) D2	Rez (ohms) D3	Rez (ohms) D4	Rez (ohms) D5	Rez (ohms) D6	Rez (ohms) D7	Rez (ohms) D8	Rez (ohms) calculated
O1 (130um)	12.2	12.1	12.6	13.6	12.35	12.2	12.15	14.12	10.46
O2 (130um)	18.6	17.3	17.8	23.8	26.2	17.4	28	22.9	16.4
O3 (130um)	14.3	13.85	15.6	17.1	14.24	13.67	14.4	18.4	12.86
I1 (55um)	12.1	11.4	11.45	13.6	12	11.51	11.8	13.77	10.32
I2 (55um)	21.2	20.84	20.6	23.6	20.15	20.63	20.34	24.05	19.27
I3 (55um)	11.2	11.3	11.63	12.9	11.53	11.4	11.74	12.6	10.16

Following the DC resistance measurements high frequency measurements were done on the special designed test structures that were included in the test wafer. As detailed in [7] the routing of the AIB bus structure between the chiplet in the 4 die package is not straight forward. In the current implementation each chiplet must communicate with two other chiplets using the AIB interface. However, the current pad arrangement of the chiplets, with the IOs pads of the AIB interface located on one side of the die creates a constrain on how to arrange the dies such equal length lines can be routed between the corresponding AIB IOs. Due to this constrain some of the lines of the bus had to include meander sections to equalize their lengths, and a maximum difference of 1.6mm in length between lines. The effects of meander lines and length difference to the signal integrity was analyzed. First frequency domain measurements were done to the test structures shown in Fig. 15. The four port measured data was then used to generate the eye diagrams using the AIB 2GBps and 900mV swing signal and terminations specified in [8]. As it can be observed in Fig. 16 and Fig. 17 the effect of the meander lines and inequal lengths has a neglectable effect on the quality of the signal integrity. The height of eye (EH) almost the same for the lines routed on layer M4, while for lines routed on layer M4 a 1.5% (12mV) degradation is noted for the lines with the meander and 1mm length difference. The eye widths (EW) are not affected at all by the difference in length and meanders. These measurement data has confirmed that the bus structures routed between the 4 chiplet functional dies will be able to support the AIB interface and it will function appropriately. These conclusions were further supported by the results published in [7], where a fully functional SIP is discussed.

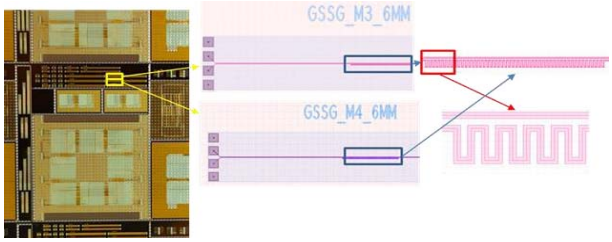


Fig. 15. Test structure to assess the signal integrity of the inequal and meander routing

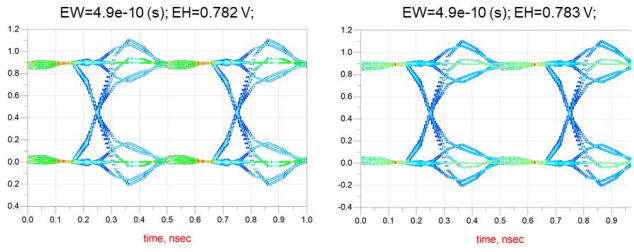


Fig. 16. Eye diagrams of 6mm coupled lines routed on M4, with equal lengths and with meander and 1mm difference in length.

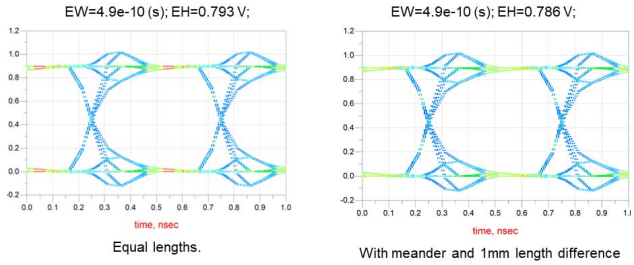


Fig. 17. Eye diagrams of 6mm coupled lines routed on M3

C. *Underfill process*

Capillary underfilling process is done after the bonding process to encapsulate the microbumps and to increase the reliability performance of the solder joints. The average filler size of the capliiary underfill material selected is 0.3um. Prior to the underfill dispensing process, the assembled wafer is subjected to a prebaking process of 125°C for 4hrs to remove any moisture followed by a plasma cleaning process for better underfill flow. As there are different die sizes and bump layout on the assembled wafer, the underfill must not overflow to the die top surface especially for the tight die to die gap for the different chiplets for the HDFOWLP MCP. Uniform flow of the underfill material is also required to encapsulate the microbumps without any underfill voids. Fig. 18 shows the underfilled MCP sample.

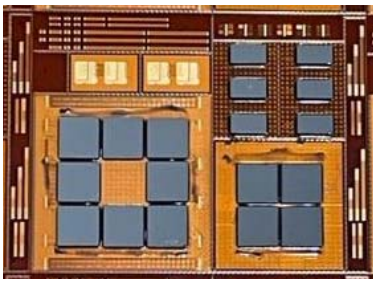
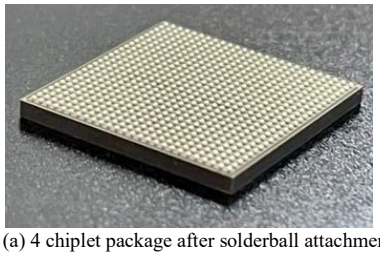


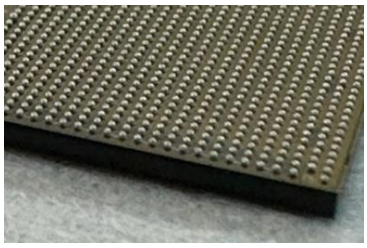
Fig. 18. Optical image on the underfilled MCP sample

D. *Solderball attachment process*

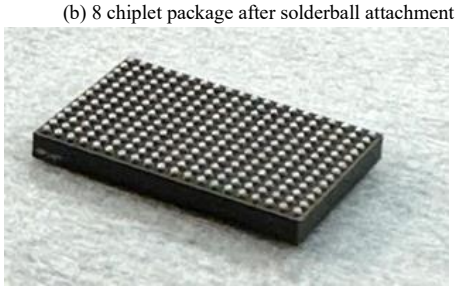
The package singulation process of the assembled wafer into 4 chiplets, 3 chiplets and 8 chiplets packaging is done after completing the molding and laser debonding process. SAC 305 solderballs of 250um diameter are attached to the Cu/Ni/Cu UBM on the singulated chiplet. The samples are then subjected to a lead-free reflow profile in the reflow oven to do the reflow process. Fig. 19 shows the images of the different chiplets after ball attachment process. Results showed no missing solder balls and no balls bridging after the ball drop process. Xray inspection also shows no solder voids.



(a) 4 chiplet package after solderball attachment



(b) 8 chiplet package after solderball attachment



(c) 3 chiplet package after solderball attachment

Fig. 19. Optical images on different MCP package after solderball attachment process

SUMMARY

We have successfully demonstrated the assembly of 3 different HDFOWLP MCP integrated with chip-to chip interconnections on the RDL-first FOWLP platform with the following observations:

a) Good bonding alignment and no solder voids are s observed on the 3 different MCP packages with high density chiplets
b) No underfill overflow on top of the test chips on the different chiplets
c) Low warpage of less than 2mm is achieved after the wafer level molding process on the RDL-first wafer
d) The MCP FOWLP package with Intel 22nm FinFet chips has passed the functional testing characterization
e) Good DC measurement data for all the 6 daisy chains and the daisy chains that connected all the 8 chiplets
f) No solder voids is achieved after the solderball attachment process

ACKNOWLEDGMENT

The University of Michigan effort is supported in part by the Defense Advanced Research Projects Agency CHIPS program and the Office of Naval Research under grant N00014-17-1-2992. This work is also supported under the grant of Neuromorphic Processor Design - Neuron Circuits (Grant ID: A1687b0033). We thank Intel for the chiplet fabrication and bumping.

REFERENCE

[1] Yoonjae Hwang, Sungwook Moon, Seungki Nam and JeongHoonAhn, "Chiplet-based System PSI Optimization for 2.5D/3D Advanced Packaging Implementation", 72nd ECTC, San Diego, pp 12 to 17

[2] Gang Duan, Yosuke Kanaoka, Robin McRee, Bai Nie, Rahul Manepalli, "Die Embedding Challenges for EMIB Advanced Packaging Technology", 71th ECTC, San Diego, pp 1 to 7

[3] Guilian Gao, Laura Mirkarimi, Gill Fountain, Dominik Suwito, Jeremy Theil Thomas Workman, Cyprian Uzoh, Bongsub Lee, KM Bang, Gabe Guevara,"Die to Wafer Hybrid Bonding for Chiplet and Heterogeneous Integration: Die Size Effects Evaluation-Small Die Applications", 72nd ECTC, San Diego, pp 1975 to 1981

[4] Chai Tai Chong, Soon Wee Ho, Chong Ser Choong, Sharon Lim PS, Surya Bhattacharya, "Heterogeneous Integration for Chiplets on FOWLP Development Line", 72nd ECTC, San Diego, pp 655 to 659

[5] Ang-Ying Lin, Yu-Min Lin, Tzu-Hsuan Ni, Tao-Chih Chang and Chang-Chun Lee, "High density batch bonding technology for chiplet design", *2022* International Symposium on VLSI Technology, Systems and Applications (*VLSI-TSA),* April 18-21, *2022*

[6] Wenchao Tian, Bin Li, Zhao Li, Hao Cui, Jing Shi, Yongkun Wang and Jingrong Zhao, "Using Chiplet Encapsulation Technology to Achieve Processing-in-Memory Functions", Micromachines 2022, 13, 1790.

[7] Teyuh Chou, Wei Tang, Mihai D. Rotaru, Chester Liu, Rahul Dutta, Sharon Lim Pei Siang, David Ho Soon Wee, Surya Bhattacharya, Zhengya Zhang, "NetfFlex: A 22nm Multi-Chiplet Perception Accelerator in High Density Fan-Out Wafer-Level Packaging", 2022 IEEE Symposium on VLSI Technology and Circuits, June 13-17, Honolulu, pp 208 to 209.

[8] Mihai D. Rotaru, Wei Tang, Dutta Rahul, Zhengya Zhang, "Design and Development of High Density Fan-Out Wafer Level Package (HD-FOWLP) for Deep Neural Network (DNN) Chiplet Accelerators using Advanced Interface Bus (AIB)", 71th ECTC, San Diego, pp 1258 to 1263