Introduction

The main package used for central processing units (CPUs) and graphics processing units (GPUs) including artificial intelligence (AI) engines is the flip chip ball grid array (FCBGA), which uses solder bumps to connect the die to the organic substrate [1, 2] (Fig. 1(a)). Owing to the increased size of organic substrates due for higher performance devices and limitations on interconnection pitch between devices and organic substrates, it is becoming increasingly difficult for conventional packaging technologies to support higher performance and speed.[3]

Fan-out wafer-level packaging (FOWLP), which eliminates the organic substrate, is a package structure in which the fine redistribution layer (RDL) formed on the LSI chip extends beyond the chip outline.[4, 5] Die-first FOWLPs have largely been adopted in mobile and consumer applications such as application processors, power management ICs, baseband processors, radio frequency ICs, and microcontrollers. However, the applicable package size is typically about 8 mm \times 8 mm or less due to the package warpage issue and cost advantages compared to conventional FCBGAs, and it has not been applied to high-performance computing.[6] The more expensive 2.5D with silicon interposer process is therefore still used, and the range of applications of FOWLP has not grown.[7, 8] To achieve small die-shift, low warpage, and multilayered fine RDL formation is mandatory to select FOWLP as a solution for high-performance computing. Compression molding using the epoxy-based liquid compression molding (LCM) material is one of the critical processes on the three challenges.[9, 10]

In previous work, we proposed a new type of 3D FOWLP...
that electrically connects the logic and memory in three dimensions by using RDL technology.[11–13] In that work, we characterized a low elasticity mold resin with photosensitivity, which eliminates the need for an expensive Cu pillar process. In the future era of higher-capacity memory, FOWLP technology is expected to find increasing demand not only for 3D structures but also for 2D structures. In this study, as in the past, we focus on the mold resin of FOWLP and propose a new 2D FOWLP structure and process to solve the above three problems (die shift, warpage, and multi RDL formation). We demonstrate a novel die-first FOWLP using a low elastic modulus resin that differs from conventional LCM. A 5-in-1 FOWLP containing a GPU die with over 2,500 pins as an AI chip and four DRAM dies placed side by side is developed in a package size of 31 mm × 31 mm (Fig. 1(b)), and the functional operation is demonstrated.

1. Issues and Solutions for FOWLP

The FOWLP fabrication process is made up of three basic flows.[14–16] Figure 2 shows a simplified image of the manufacturing process flows: The electrical connections between the die and package are implemented by direct RDL formation on the die pad in a die-first (face-down) process;[17, 18] RDL formation is performed after CMP using a die-first (face-up) process;[19, 20] and flip chip mounting to RDL is performed using solder bumps on the chip in an RDL-first process.[21, 22]

In this study, a die-first (face-down) process is selected as the manufacturing process because the corresponding mass production technology has already been established. Unlike other processes, the die-first (face-down) process does not require any components (e.g., Cu pillars, solder bumps) to connect to the RDL, which reduces the manufacturing cost.

Our objective is to realize an unprecedented module that integrates five chips and forms a multilayer RDL in a 31 mm × 31 mm FOWLP. The following are reported as major technical issues for current FOWLPs.

1. Die shift
2. Warpage of reconfigured wafer
3. Multi-layered fine RDL formation

Die shift describes movement of the chip after placement during resin molding, debonding of the carrier, and cooling of the reconfigured wafer.[23–25] Warpage of the reconfigured wafer is caused by the difference in thermal expansion coefficient between the mold resin and the carrier substrate (Si or glass or organic, etc.).[26–28] These issues are mostly affected by the mechanical properties of the mold resin. In order to solve this issue, a low elastic modulus material is adopted for the mold resin that forms the reconfigured wafer. In addition, a new process for controlling warpage by placing the Si support substrate on the backside of the mold resin is developed. Using the reconfigured wafer reduced die shift and wafer warpage, the process of implementing a multi-layered fine RDL to create a 5-in-1 FOWLP is evaluated.

2. 5-in-1 FOWLP Technology

2.1 Layer-pattern designs

In order to make the wiring length between the GPU and each DRAM the same, the four DRAM chips were placed around the GPU chip. Signal integrity (SI) simulation and power integrity (PI) simulation were performed using the designed pattern. Frequencies of up to 800 MHz were used in the SI simulation, and voltage and current values of up to 1.5 V and 20 A were used in the PI simula-
tion. Although the initial design used two RDL layers, the results of the SI simulation found that the eye diagram did not meet the criterion (±0.3 V) for some clocks (Fig. 3(a)).

The design was therefore updated to use three RDL layers with an enhanced power supply in the second layer. As a result of the redesign, SI simulation of the three RDL layers showed a good eye diagram (Fig. 3(b)). Figure 4 shows the wiring layout with three RDLs of FOWLP. The first layer is mainly signal and power, the second layer is mainly power, and the third layer is mainly power and ground.

2.2 Mold material

In order to prevent die shift and achieve low warpage, a mold resin with a low elastic modulus that has excellent stress relaxation was employed. Table 1 shows the mechanical properties of the developed mold resin. The base of the material is silicone and does not contain any filler. The Young’s modulus of the material is about 50 MPa, which is very low compared with the materials used in conventional FOWLPs (>20 GPa).[29, 30]

2.3 Fabrication process

Table 2 shows the GPU and DRAM die specifications. The GPU has a chip size of 14.3 mm × 8.9 mm, a pad pitch of 180 μm, and 2,500 pads, and the DRAM has a chip size of 8.1 mm × 9.4 mm, a minimum pad pitch of 82 μm, and 155 pads. Table 3 shows the package specifications for the 5-in-1 FOWLP. The package was designed to be the same size and have the same number of external pins as the GPU package in the computer circuit board, as shown in Fig. 1(a), in order to demonstrate the operation of the system board. The process flow of the developed FOWLP is shown in Fig. 5. First, after the release layer was deposited on an 8-inch diameter glass substrate (coefficient of thermal expansion [CTE] 3.4 ppm/K), the die was mounted facedown using a flip chip bonder (FCB3, Panasonic). During the mounting process, the fiducial marks formed on the glass substrate were recognized and the die was moved to a pitch of 31.08 mm. The mounting conditions

---

**Table 1**  Material properties of mold compound.

| Item          | Unit | General properties     |
|---------------|------|------------------------|
| Material type | –    | Silicone resin w/o filler |
| Appearance    | Color | Black                   |
| Young’s modulus | MPa | about 50                |
| Elongation    | %    | about 50                |
| CTE           | ppm/K| about 170               |

**Table 2**  Die specifications.

| GPU          | Size (mm) | 14.3 × 8.9 |
|--------------|-----------|------------|
| Pad pitch (μm) | 180 (min) |            |
| Pad count    | >2,500    |            |
| Thickness (μm) | 150      |            |
| DRAM         | Size (mm) | 8.1 × 9.4  |
| Pad pitch (μm) | 82 (min) |            |
| Pad count    | 155       |            |
| Thickness (μm) | 150      |            |

**Table 3**  Package specifications.

| Size (mm) | 31 × 31 |
|-----------|---------|
| BGA ball pitch (mm) | 1        |
| BGA pad count | 900 (30 × 30) |
| BGA ball height (mm) | 0.5      |
| RDL         | Number of layers | 3         |
|             | Metal 1            | L/S = 10 μm/20 μm |
|             | Metal 2            | L/S = 20 μm/20 μm |
|             | Metal 3            | L/S = 30 μm/30 μm |
|             | Via size (μm)      | 35μ      |
|             | Land size (μm)     | 55μ       |
used were a 9 N load, 60°C, and 500 ms. The mold resin was then deposited using a printing method. A metal mask was placed on the top surface of the glass substrate on which the die was mounted, and two printings were performed using conditions of 4 kgf load and 5 mm/s squeegee speed at room temperature. The thickness of the mold resin on the backside of the embedded chips was adjusted to be about 20 µm. After printing, the mold resin was temporarily cured (100°C/30 min). Although the mold resin has a low elastic modulus and is expected to prevent die shift and reduce warpage, the mold resin cannot stand on its own and has issues concerning transportability during manufacturing. Therefore, after forming the mold resin, an 8-inch-diameter Si support substrate (725 µm thickness) was bonded to the back of the mold resin at the wafer level using wafer bonder. The bonding conditions used were 8 kN load, 100°C, and 180 s. Next, the mold resin was subjected to main curing (210°C/4 h). The Si support substrate was used as part of the product. Next, the glass substrate and mold resin were peeled off at the interface of the release layer to expose the electrodes of the die. RDLs with three layers of Cu plating were then formed using a semi-additive process. Siloxane-based dielectric material with material properties similar to mold resin was used as the insulating material, and a via of minimum diameter 35 µm was formed on land of diameter 55 µm using an exposure dose of 1,600 mJ/cm². Ti (45 nm)/Cu (300 nm) was formed as the seed layer using a sputtering system, and the RDL was formed using a minimum line and space rules of 10 µm/20 µm. In the three-layer RDL, the thickness of the Cu wiring was 8 µm and the thickness of the dielectric layer was 15 µm. Finally, after mounting BGA balls, the 5-in-1 FOWLP was fabricated in 31-mm square pieces.

3. Results and Discussion

3.1 Observation of RDLs

Figure 6 shows the BGA side and the backside of the FOWLP. Pitch BGA balls of 1 mm were mounted on the package surface side, and a 725-µm-thick Si substrate was placed on the backside of the mold resin. Figure 7 shows a top view of the RDLs, showing the 1st metal (line width 10 µm), 2nd metal (line width 20 µm), 3rd metal (line width 30 µm), and BGA. Figure 8 shows a cross-section of the package, where the Si support substrate is placed on the backside of the GPU die through the mold resin, and the three-layer RDL with total thickness of 55 µm can be seen from the pads of the GPU die.

3.2 Die shift measurement

The die shift was evaluated by fabricating 21 packages of size 31 mm × 31 mm using an 8-inch wafer glass support substrate. Figure 9(a) shows a photograph of the GPU and
DRAM after mounting on the glass substrate, and Fig. 9(b) shows a photograph after the glass substrate was removed. The amount of die shift was measured at the pad position in the GPU in the package after peeling off the glass substrate. Figure 10 shows the amount of die shift of the GPU at five locations (center, top, right, bottom, left) in the package. This shows that the amount of die shift was less than ±5 μm, which is within the criterion of ±10 μm, indicating that the amount of die shift in the study is very small. This shows that employing a low elasticity mold resin is effective for reducing the die shift in 8-inch wafers.

### 3.3 Warpage measurement

The warpage of the package was evaluated at the package level and the mounted board level, respectively. The Si surface on the backside of the package was measured using a 3D laser measuring instrument. Figure 11(a) shows the result of the package level warpage measurement. The amount of warpage of the package was 10.6 μm, which is very small for a package size of 31 mm × 31 mm.

It is thought that this is thanks to the effect of the rigid Si support substrate placed on the backside of the low elasticity mold resin. Next, the package was placed on a computer circuit board and connected to the board using a reflow soldering processes with 900 BGA balls, as shown in Fig. 12(a). The warpage of the mounted package was...
measured in the same way. The warpage of the package was 125.5 \( \mu m \), which is an increase compared with the package level, as shown in Fig. 11(b). However, the amount of warpage was small compared to the pitch of BGA balls (1 mm) and BGA ball height (0.5 mm), and the connection of all 900 pins was confirmed.

### 3.4 Operation Verification

The sub-board on which the prototype 5-in-1 FOWLP was mounted was connected to a console equipped with input and output functions, as shown in Fig. 12(b). A 12-V power supply was then supplied to verify the operation of the FOWLP. As a result, the GPU was confirmed to operate at an operating frequency of 800 MHz through the three-layer RDL of the FOWLP at a voltage of 1.0 V.

### 3.5 Thermal stress analysis

Stress simulation analysis was performed using a new siloxane-based mold resin and inter layer dielectrics with low elasticity and high CTE and a standard material currently used in the mass production of FOWLP. Table 4 shows the specific properties of the standard material. The simulation was analyzed at the package level. ADVENTURECluster™ finite element analysis software was used for the stress simulation. In the simulation, all of the materials behaved as isotropic linear elastic materials. In order to compare the relative stresses of the materials at different curing temperatures, the effective stress-free temperature was taken to be room temperature. In the simulation, the equivalent stress on the via (35 \( \mu m \) diameter, 10 \( \mu m \) height) on the embedded die was compared under two temperature conditions (−55°C and 125°C).

The simulation results showed that the maximum stress at −55°C for the new material was 591 MPa at the top of the via, while that at 125°C was 471 MPa at the bottom of the via, indicating that the stress concentration point changes with temperature (Fig. 13). In the standard material, the maximum stress occurred at the bottom of the via, independent of the temperature condition. Figure 14 shows the stress analysis results for the new material and the standard material at −55°C and 125°C. The RDL between vias of the standard material was flat, while that of the new material was deformed under each temperature condition. In addition, the maximum stress of the new material was about 2.8 times higher at −55°C and about 1.7 times higher at 125°C than that of the standard material, as shown in Fig. 15. From these results, it was confirmed that the new material had a tendency to increase the maximum stress owing to deformation of the wiring caused by thermal shrinkage because of the high CTE. In future develop-

---

**Table 4** Material properties of the standard materials.

| Item                  | Unit | General properties         |
|-----------------------|------|-----------------------------|
| **Mold material**     |      |                             |
| Material type         | –    | Epoxy resin                 |
| Young’s modulus       | GPa  | 22                          |
| Tg                    | °C   | 165                         |
| CTE                   | ppm/K| 7                           |
| **Inter layer dielectrics** | |                             |
| Material type         | –    | Photosensitive phenolic resin |
| Young’s modulus       | GPa  | 3.4                         |
| Tg                    | °C   | 300                         |
| CTE                   | ppm/K| 46                          |
ment, it will be necessary to optimize the material property values of the siloxane-based dielectric material, such as low CTE and high elasticity, to ensure the package reliability.

This study attempted to solve issues with FOWLP such as die shift and warpage by introducing new materials into the die-first (face-down) process. We expect new value to be created through cooperation among design, device, substrate, assembly, material, and equipment manufacturers for future panel-level packages.

4. Conclusions
A 5-in-1 FOWLP consisting of a GPU die with over 2,500 pins and four DRAM dies placed side by side was developed by adopting a low elastic modulus mold resin and bonding a Si substrate to the backside of the reconfigured wafer. Integrating multiple chips into a single package with a novel material is expected to significantly reduce the mounting area, provide high-speed transmission, and reduce power consumption. The developed package achieved a die shift of less than ±5 µm, package warpage of 10.6 µm over 31 mm × 31 mm, and three-layer RDL formation with L/S = 10 µm/20 µm. Function tests were demonstrated through a three-layer RDL of the FOWLP. This paper shows the advantages and feasibility of FOWLP for AI edge computing and large memory modules.

Outstanding themes/concerns to be addressed as a next step for determining the possibility of using FOWLP for AI edge computing are as follows:

1) Development of siloxane-based resin with optimized CTE and elasticity to manage both die-shift and package reliability.
2) Implementing reliability tests using actual samples to confirm simulation results.
3) Quantitative confirmation the improvement in system performance by adopting a 5-in-1 FOWLP.

Acknowledgements
This study is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO) to develop cross-sectoral technologies for IoT promotion. The authors wish to acknowledge Shin-Etsu Chemical Co., Ltd. for material and process support. The authors thank Shuzo Akejima, Hirokazu Ezawa and Takafumi Fukuda for their encouragement and useful discussions on this research.

Company names, product names, and service names may be trademarks of their respective companies.

References
[1] C.-T. Peng et al., “Reliability Analysis and Design for the Fine-Pitch Flip Chip BGA Packaging,” IEEE TRANSACTIONS ON COMPONENTS AND PACKAGING TECHNOLOGIES, Vol. 27, No. 4, pp. 684–693, DECEMBER 2004.
[2] Y. Jung et al., “Development of Large Die Fine Pitch Flip Chip BGA using TCNCP Technology,” 2012 IEEE 62nd Electronic Components and Technology Conference, p. 439, 2012 May.
[3] W. S. Tsai et al., “Generational changes of flip chip interconnection technology,” 12th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), pp. 306–310,
[4] T. Meyer et al., “Embedded Wafer Level Ball Grid Array (eWLB),” 2008 10th Electronics Packaging Technology Conference, p. 994, 2008 Dec 9.

[5] Y. Jin et al., “Next generation eWLB (embedded wafer level BGA) packaging,” 2010 12th Electronics Packaging Technology Conference, p. 520, 2010 Dec 8.

[6] A. P. Lujan, “Cost Comparison of Fan-out Wafer Level Packaging and Flip Chip Packaging,” IMAPS 2017, pp. 1–5, 2017.

[7] S. Y. Hou et al., “Wafer-Level Integration of an Advanced Logic-Memory System Through the Second-Generation CoWoS Technology,” IEEE TRANSACTIONS ON ELECTRON DEVICES, Vol. 64, No. 10, p. 4071, OCTOBER 2017.

[8] L. Lin et al., “Reliability characterization of Chip-on-Wafer-on-Substrate (CoWoS) 3D IC integration technology,” 2013 IEEE 63rd Electronic Components and Technology Conference, p. 366, 2013 May.

[9] K. Ueno et al., “Development of liquid, granule and sheet type epoxy molding compounds for fan out wafer/panel level package,” 2016 11th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), p. 285, 2016 Oct.

[10] T. Braun et al., “Large area compression molding for Fan-out Panel Level Packaging,” 2015 IEEE 65th Electronic Components and Technology Conference (ECTC), p. 1077, 2015 May.

[11] K. Mori et al., “3D Fan-Out Package Technology with Photosensitive Through Mold Interconnects,” 2019 IEEE 69th Electronic Components and Technology Conference (ECTC), pp. 1140–1145, 2019.

[12] K. Mori et al., “3D Packaging and Integration Technology using Photosensitive Mold,” 2020 International Conference on Solid State Devices and Materials (SSDM), pp. 131–132, 2020.

[13] K. Mori et al., “3D integration technology with photosensitive mold for fan-out package,” Japanese Journal of Applied Physics (JJAP), Vol. 60, SBBC03, 2021.

[14] T. Braun et al., “Panel Level Packaging - A View Along the Process Chain,” 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), p. 70, 2018 May.

[15] M. Shih et al., “Electrical, Thermal, and Mechanical Characterization of eWLB, Fully Molded Fan-Out Package, and Fan-Out Chip Last Package,” IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY, Vol. 9, No. 9, pp. 1765–1773, SEPTEMBER 2019.

[16] M. Woehrmann et al., “Ultra-thin 50 um Fan-Out Wafer Level Package: Development of an innovative assembly and de-bonding concept,” 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), p. 675, 2018 May.

[17] J. Lau et al., “Feasibility Study of Fan-Out Wafer-Level Packaging for Heterogeneous Integrations,” 2019 IEEE 69th Electronic Components and Technology Conference (ECTC), p. 903, 2019 May.

[18] J. H. Lau et al., “Fan-Out Wafer-Level Packaging for Heterogeneous Integration,” IEEE Transactions on Components, Packaging and Manufacturing Technology, Vol. 8, Issue 9, p. 1544, Sept. 2018.

[19] J. H. Lau et al., “Design, Materials, Process, Fabrication, and Reliability of Fan-Out Wafer-Level Packaging,” IEEE Transactions on Components, Packaging and Manufacturing Technology, Vol. 8, Issue 6, p. 991, June 2018.

[20] X. Hua et al., “Development of chip-first and die-up fan-out wafer level packaging,” 2017 IEEE 19th Electronics Packaging Technology Conference (EPTC), p. 1, 2017 Dec.

[21] H.-C. Lu et al., “Chip Last Fan-Out Packaging for Millimeter Wave Application,” 2016 IEEE 66th Electronic Components and Technology Conference (ECTC), p. 1303, 2016 May.

[22] Y.-M. Lin et al., “An RDL-First Fan-out Wafer Level Package for Heterogeneous Integration Applications,” 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), p. 349, 2018 May.

[23] Y. Han et al., “Comprehensive Investigation of Die Shift in Compression Molding Process for 12 Inch Fan-Out Wafer Level Packaging,” 2016 IEEE 66th Electronic Components and Technology Conference (ECTC), p. 1605, 2016 May.

[24] T. Braun et al., “24”x18” Fan-out panel level packing,” 2014 IEEE 64th Electronic Components and Technology Conference (ECTC), p. 940, 2014 May.

[25] A. Podpod et al., “A Novel Fan-Out Concept for Ultra-High Chip-to-Chip Interconnect Density with 20-µm Pitch,” 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), p. 370, 2018 May.

[26] E. K. Th et al., “Encapsulation challenges for wafer level packaging,” 2009 11th Electronics Packaging Technology Conference, p. 903, 2009 Dec.
[27] H.-W. Liu et al., “Warpage characterization of panel Fan-out (P-FO) package,” 2014 IEEE 64th Electronic Components and Technology Conference (ECTC), p. 1750, 2014 May.

[28] J. H. Lau et al., “Warpage Measurements and Characterizations of Fan-Out Wafer-Level Packaging With Large Chips and Multiple Redistribution Layers,” IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY, Vol. 8, No. 10, p. 1729, OCTOBER 2018.

[29] K. Kan et al., “The Novel Liquid Molding Compound for Fan-out Wafer Level Package,” 2016 International Conference on Electronics Packaging (ICEP), pp. 557–561, 2016.

[30] T. Kamimura et al., “Development of Liquid Compression Molding (LCM) Material for Low Warpage,” 50th International Symposium on Microelectronics, pp. 25–28, 2017.

Kentaro Mori
Soichi Yamashita
Masahiro Sekiguchi