IMPLEMENTATION OF A FULLY MOLDED FAN-OUT PACKAGING TECHNOLOGY

B. Rogers, C. Scanlan, and T. Olson
Deca Technologies, Inc.
Tempe, AZ USA
boyd.rogers@decatechnologies.com

ABSTRACT
Fan-Out Wafer-Level Packaging (FOWLP) is finally beginning to gain traction as a means of packaging semiconductor chips with interconnect densities exceeding the capabilities of standard Wafer Level Chip Scale Packaging (WLCSP). However, despite its promise, widespread adoption of FOWLP has not yet occurred − due largely to reported manufacturablity, yield and reliability concerns as well as cost-effectiveness as compared with existing laminate BGA packaging.

This paper describes a novel fully molded FOWLP technology which addresses the noted concerns. In this approach, Cu pillars are fabricated at IO locations on the native semiconductor wafer prior to the panelization process. Then, during panelization, the front and sides of the die are fully covered with mold compound, with the Cu pillars providing current pathways through the mold compound on the front die surface. Next, buildup processing is performed on the front panel surface with interconnection between the exposed Cu pillars and a fan-out bump array. Finally, an epoxy laminate is applied to the backside of the panel to fully encase the die in epoxy. The benefits of the fully molded structure include: (1) elimination of the discontinuity at the die surface between the edge of the die and the mold compound; and (2) improved board level reliability as a result of the molded layer separation between the chip and the printed circuit board (PCB) connections. The end result is a rugged package.

The fully molded package has been coupled with an Adaptive Patterning® methodology which compensates for die shifts introduced in pick and place, molding and curing operations during panelization. The dynamic routing produced by Adaptive Patterning allows the design on each package on every panel to be customized for the actual die location. This enables higher pick and place speeds, tighter via ground rules, higher yield, and lower overall costs.

The two approaches together represent a significant departure from past fan-out technologies and help address many of the manufacturability, cost, yield, and reliability issues that have limited widespread adoption of this technology. This paper details the construction of a fully molded FOWLP built using Adaptive Patterning and provides modeling and reliability data validating the approaches.

Key words: Adaptive Patterning, fan-out wafer-level packaging (FOWLP), panelized packaging, WLP

INTRODUCTION
The handheld consumer electronics space, where portability and increasing functionality are strong drivers, continues to motivate the transition to packaging approaches that provide small size, high performance, and low cost. Wafer Level Chip Scale Packaging (WLCSP), which offers the smallest packaging form factor, has often been a preferred option for addressing the handheld market. In WLCSP, chip IOs are generally fanned-in across the die surface using polymer and redistribution line (RDL) buildup layers to produce an area array, and large solder bumps are formed at the terminals by ball drop or plating. These additive processes allow the chip to be attached directly to a PCB with high reliability. However, two progressions in front-end chip manufacturing bring challenges to packaging in a WLCSP format: (1) die shrink, enabled by advancing semiconductor technology nodes, makes it increasingly difficult to fit all of the large solder ball IOs on the die surface; and (2) increasing chip functionality produces a need for more IOs, also making WLCSP packaging more difficult. One approach to extending WLCSP is to shrink the size of the IOs or solder bumps on the chip surface so that more can fit within the chip area. However, this approach is generally limited by a lack of assembly infrastructure, design rule limitations in end application PCBs or higher assembly costs in managing the smaller or tighter pitch IOs. [1]

Fan-out, or FOWLP has been offered for a number of years as an alternative for addressing constraints to WLCSP [2,3]. In this technology, chips are singulated and then embedded in a panel. A common method for forming this panel is to place the chips face-down on a carrier at a desired pitch and then mold over them using compression or print molding. The molded panel is subsequently separated from the carrier. The panel is often formed in the shape of a wafer, so that standard wafer processing techniques can be used to create buildup layers on the panel surface. The extra panel surface around the chip allows IOs to be both fanned in over the chip and fanned out across the mold compound, thus

accommodating a larger number of IOs. After buildup layer processing and solder ball attachment, the packages undergo backgrinding, laser marking, and singulation, just like WLCSPs. The resulting package is often just slightly larger than the chip and just large enough to accommodate the IOs. Like WLCSPs, the package is ready to be mounted directly to an end application PCB.

The potential benefits of FOWLP are numerous. This technology provides the smallest and lightest possible form factor for packaging small, high IO chips that cannot be packaged as WLCSPs. As with WLCSPs, the device-to-board connections through thick copper routing layers and large solder balls offer excellent electrical properties and performance. When hitting acceptable cost targets, FOWLP can potentially displace other forms of packaging, such as flip chip or wirebond BGAs. In those cases, it generally brings a size advantage and eliminates the need for custom substrates, significantly simplifying the supply chain. Finally, FOWLP enables the connection of two or more chips in the fan-out routing layer, facilitating multichip and system-in-package (SIP) applications [4-7].

Cost, yield and reliability issues have effectively limited the widespread adoption of FOWLP despite its promise. Placing singulated chips on the carrier to form the molded panel requires high placement accuracy. Any misplacements can lead to pattern overlay difficulties in the buildup process on the reconstituted panel. The requirement for high placement accuracy restricts throughput at the pick-and-place operation, leading to high process costs. During the molding operation and mold cure, die drift or movement can occur. This die drift can further complicate pattern overlay matching in the buildup process on the panel and can result in yield loss when the drift is excessive. Addressing or overcoming die offset and resulting overlay issues is one of the keys to making FOWLP competitive with other package formats.

Other challenges are posed by performing the fanout routing over a surface which is partially composed of silicon and partially of mold compound. These challenges are illustrated in Figure 1 for a conventional FOWLP structure. During the molding process, mold can seep in under the edge of the face-down die. If this mold flash extends far enough, it can cover bond pads and result in yield loss. The discontinuity posed by the transition between the silicon chip and the mold compound at the die surface can result in a severe topography step which is difficult to route over with the RDL, even with substantial smoothing of a base polymer layer. Finally, the CTE mismatch between the silicon and the mold compound can result in stress in the fanout RDL structure and potential fracture of the RDL during reliability stressing.

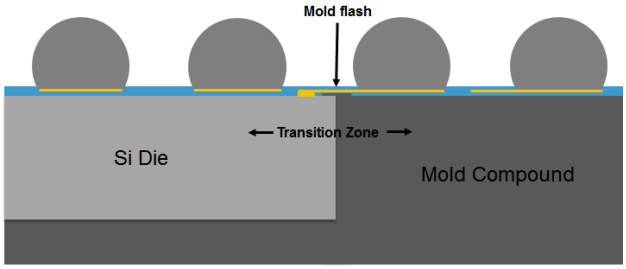


Figure 1. Illustration of a conventional FOWLP structure, showing the transition zone between silicon and mold compound on the package front surface

This paper discusses two approaches for addressing several of the key issues associated with conventional FOWLP technology. By creating a fully molded structure, where mold compound protects the front surface of the die, the issues cited above associated with the silicon – mold transition are eliminated , resulting in a more robust package. A second enabling technology, called Adaptive Patterning, allows die offset to increase by an order of magnitude compared with conventional methods and can facilitate tighter ground rules, higher yield, and lower overall costs. Taken together, the two approaches offer strong potential to overcome basic issues associated with FOWLP technology and enable wide spread adoption.

FULLY MOLDED FOWLP STRUCTURE

A fully molded FOWLP structure is illustrated in Figure 2. In this approach, Cu pillars are fabricated at IO locations on the native semiconductor wafer prior to the panelization process. Then, during panelization, the front and sides of the die are covered with mold compound, with the Cu pillars providing current pathways through the mold on the front die surface. Next, buildup processing is performed on the front panel surface with interconnection between the exposed Cu pillars and a bump array. Finally, an epoxy laminate is applied to the backside of the panel to fully encase the die in epoxy.

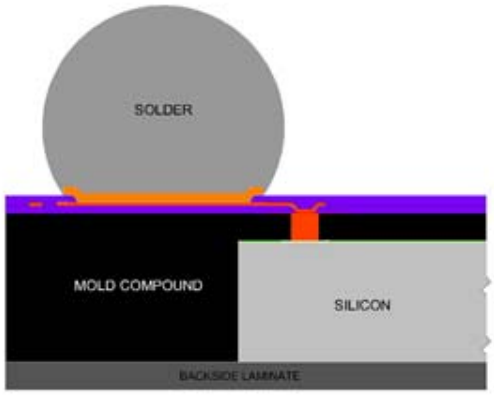


Figure 2. Illustration of fully molded FOWLP structure

The process flow for the fully molded FOWLP is shown in Figure 3. In the first segment of processing, termed ‘Wafer Prep’, Cu pillars are fabricated on the native wafer

by sputter-depositing seed layers, patterning a thick photoresist, plating Cu pillars, and then performing strip and etch of the photoresist and seed layers. In the second 'Panelization' segment, the chips are singulated and attached to a carrier face up. Overmolding is applied, and the newly formed panel containing the chips is debonded from the carrier. The panel is then front ground to reveal the Cu pillars. Finally, an optical scanner is used to measure the location of each die on the panel, to enable the Adaptive Patterning technique described in the next section. Following 'Panelization', 'Fan-out' processing is performed on the molded panel. This includes patterning and curing a polymer layer, patterning and plating Cu RDL, processing a second polymer layer, and, optionally, patterning and plating UBM pads. Ball drop and reflow are then performed. Finally, panel backgrind is performed and a backside laminate is applied, followed by laser mark, package saw, and tape and reel.

An example structure built using a fully molded FOWLP process flow is shown in Figure 4. The chip has been completely encased in epoxy, forming a robust package, and the discontinuity at the die edge which exists on conventional FOWLP structures has been eliminated.

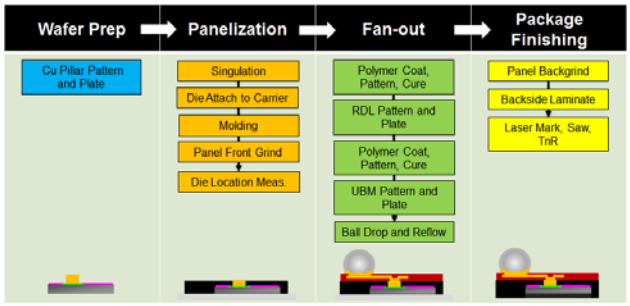


Figure 3. Fully molded FOWLP process flow

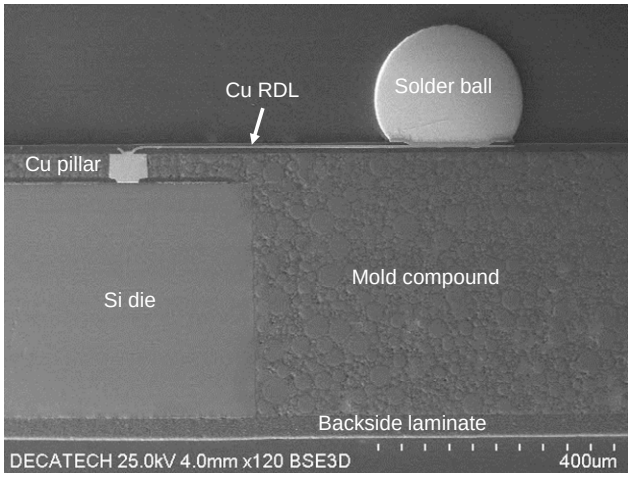


Figure 4. Cross-section of a fabricated fully molded FOWLP

ADAPTIVE PATTERNING

A method called Adaptive Patterning has been developed to correct for die shifts inherent in chips-first, panelized packaging processes [8]. For the fully molded FOWLP structures described here, the Adaptive Patterning process works by dynamically adjusting a portion of the interconnect structure to accurately connect to the Cu pillars protruding through the mold compound for each individual die in the molded panel. The flow for Adaptive Patterning is summarized in Figure 5. At the end of the 'Panelization' process, an optical scanner is used to inspect the Cu pillar protruding through the mold compound for each die, to determine the actual position and rotation of every die on the panel with respect to the ideal design frame of reference. A proprietary design tool adjusts the fan-out unit design for each package on the panel so that the first via layer and fan-out RDL pattern are properly aligned to the pillars on the die. The design files for each panel are imported to a lithography machine which uses the design data to dynamically apply a custom, Adaptive Pattern to each panel.

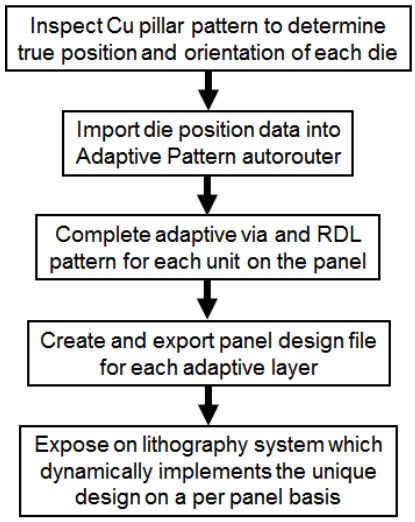


Figure 5. Adaptive Patterning flow

The process for creating the Adaptively Patterned via and fan-out RDL layers is further illustrated in Figure 6. First, a nominal fan-out RDL design is created. Then, a partially routed RDL layer called a prestratum is formed by omitting a small portion of the RDL layer in close proximity to the Cu pillars. After scanning the panel to measure the actual position and orientation of each unit, the design of each unit on the panel is completed to connect the prestratum pattern to the Cu pillar pads and their corresponding dielectric vias. The adaptive region in which the RDL traces are allowed to dynamically change is typically on the order of 100um to 200um.

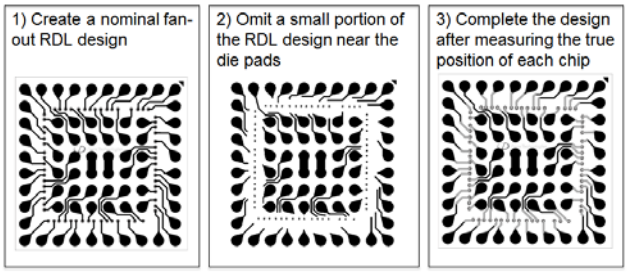


Figure 6. Process for creating Adaptively Patterned via and fan-out RDL layers

TEST VEHICLE AND PACKAGE RELIABILITY

One of the test vehicles being used to validate the fully molded FOWLP approach is shown in Figure 7. The die is a 3.0×3.0mm^2 test chip, with aluminum pads strategically connected to form daisy chains in the completed package. The fully molded FOWLP is 4.1×4.1mm^2, with a body thickness of 380um. The finished package contains 72 balls on a 0.4mm pitch. Nets are available for open-short testing or for completing the daisy chains on the board side, allowing for real time monitoring during board level reliability testing.

Figure 7. FOWLP test vehicle: 4.1×4.1mm^2, with 72 balls on a 0.4mm pitch

The test vehicle after 'Panelization' and prior to 'Fan-out' routing is shown in Figure 8. In a macroscopic view, the Cu pillars can be seen protruding through the mold compound but the die are also partially visible through the thin mold compound layer. In the microscopic view, only the pillars and the corner fiducials are clearly visible.

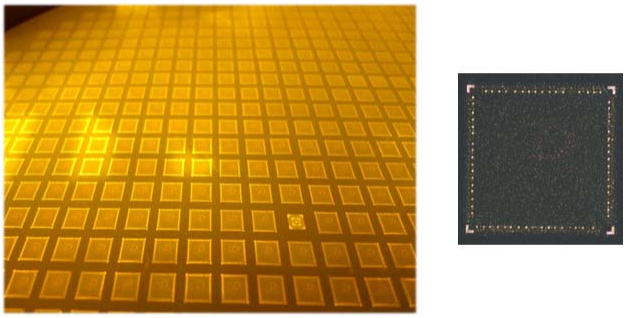


Figure 8. Test vehicle after 'Panelization' and prior to 'Fan-out' routing

The completed fully routed package is shown in Figure 9.

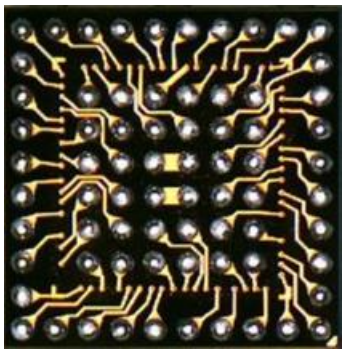
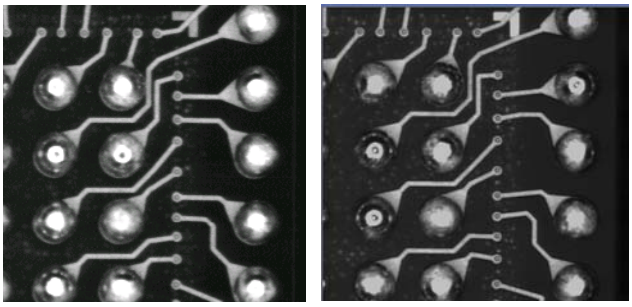


Figure 9. Routed, fully molded FOWLP

Figure 10 illustrates the effectiveness of Adaptive Patterning in completing the routing on this device. In Figure 10(a), the routing in the upper right hand corner of the package is shown, for a die that was offset from the original design position by -4.4um in X, +5.7um in Y, and -0.01 rotation error. Figure 10(b) shows the same corner for a die offset by +8.2um in X, -21um in Y, and +0.13 degrees. The routing near the Cu pillars has been dynamically adjusted to account for the die shifts, resulting in good alignment in both cases.



(a) **(b)**

Figure 10. Effectiveness of Adaptive Patterning in correcting for die offset: (a) offset from design position by -4.4um in X, +5.7um in Y, and -0.01 degrees in rotation; (b) offset by +8.2um in X, -21um in Y, and +0.13 degrees in rotation.

Package level reliability testing has been completed on this part, and the results are shown in Figure 11. The part has passed HTS, unbiased HAST, and Temperature Cycling, with MSL3 preconditioning. Board level reliability testing is currently underway.

Stress Test	Test Condition	Failure Criteria	Sample Size	Result
Cross section	1 unit per Test Vehicle	Visual and measurements	1	Passed
External Visual	5 units per Test Vehicle	Visual	5	Passed
Ball Shear, Time Zero	10 shears, 10 packages	Shear strength	10	Passed
Electrical Test, Time Zero	Open and Short testing, 35 nets	Electrical	231	Passed
Highly Accelerated Stress Test	130ºC, no bias, 96h Precondition: J-STD-020 Moisture Sensitivity Level 3, 24hrs, 85ºC, 85%RH + 3X Reflow, 260ºC	Electrical and visual	77	Passed
Temperature Cycle Test	JESD22-A104, Condition B, -55ºC to +125ºC, 500cyc Precondition: J-STD-020 Moisture Sensitivity Level 3, 24hrs, 85ºC, 85%RH + 3X Reflow, 260ºC	Electrical and visual	77	Passed
High Temperature Storage	150ºC, 500 hrs	Electrical and visual	77	Passed

Figure 11. Package qualification results for 4X4mm^2 0.4mm pitch 72 IO test vehicle

MODELING WORK

Modeling work was performed to predict the thermal cycling performance of the fully molded FOWLP. The test vehicle assumed in the modeling work is shown in Figure 12. This was an 8X8mm^2 FOWLP with 324 IOs on a 0.4mm pitch. The packages were assumed to be mounted to 1mm thick printed circuit boards (PCBs) with non-solder-mask defined PCB pads. Standard JEDEC conditions were assumed for the temperature cycling (method G: -40 to 125°C, 1 cycle per hour) [9].

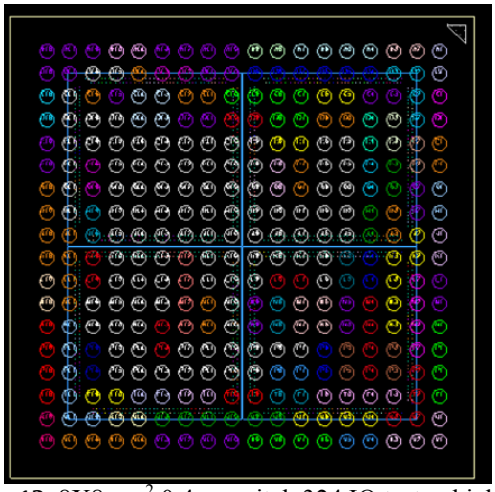


Figure 12. 8X8mm^2 0.4mm pitch 324 IO test vehicle

An ANSYS model, illustrated in Figure 13, was constructed to simulate the thermal cycling performance of the test vehicle. Symmetry was used to reduce the model to ¼ of the package size. For thermal cycling, the critical joint is at the corner bump, which is the furthest bump location from the neutral point, the package center. The strain energy-density-distribution (SED) for the corner bump can be used to predict the thermal cycle lifetime of the part [10,11]. By determining the SED accumulated per cycle, a value for the thermal cycle lifetime was derived.

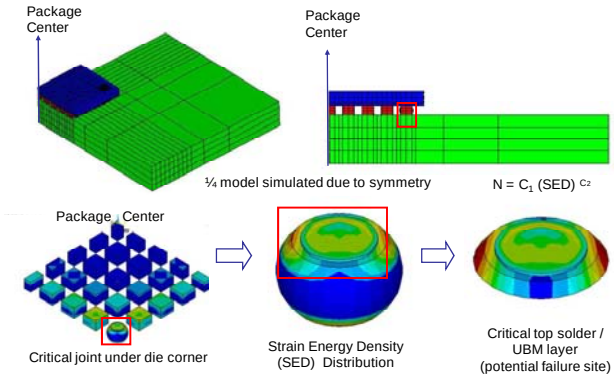


Figure 13. ANSYS model used to simulate 8X8mm^2, 0.4 mm pitch, 324 ball fully molded package

Results of the modeling work are shown in Figure 14, where the strain energy density of the corner bump on the package side is shown. Despite the large size of the package, the modeling predicts a first failure of 970 cycles. The mold compound on the front of the die in the fully molded FOWLP increases the mean coefficient of thermal expansion (CTE) of the package, thus reducing the CTE mismatch between the component and PCB and increasing the thermal cycling life.

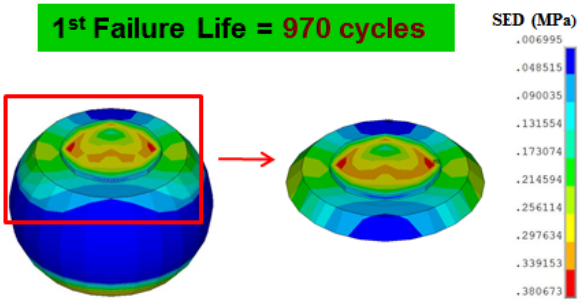


Figure 14. Strain energy density of the corner bump on the package side and the predicted thermal cycling life for the 8X8mm^2 fully molded package.

CONCLUSIONS

Two new methods have been described for Fan-Out Wafer-Level Packaging: (1) a fully molded structure, with mold compound covering the active surface of the die and eliminating the die edge discontinuity found in the conventional approach; and (2) Adaptive Patterning which electronically compensates for the die shifts inherent in FOWLP technology. The two approaches can together address the key manufacturability, yield, and reliability issues that have limited widespread adoption of FOWLP technology. Additionally, running M-Series panels through Deca's unique wafer level foundry enables a cost-effective flow. The modeling work suggests that a fully molded FOWLP can provide good thermal cycling life, even at large package sizes. Actual board level

cycling and drop tests will be the subject of follow-on work.

ACKNOWLEDGMENT

The authors would like to acknowledge SMARTs Enterprise LLP for its assistance in performing the simulation work in this study. The authors also would like to thank members of the Deca team for their help in collecting the data for this study.

*Adaptive Patterning™ by Deca Technologies, Inc.

REFERENCES

[1] Anderson, R., et. al. "Advances in WLCSP Technologies for Growing Market Needs," *IWLPC Proceedings*, Oct. 2009.

[2] Brunnbauer, M. et. al., "Embedded Wafer Level Ball Grid Array (eWLB)," *Electronics Packaging Technology Conference 8th Proceedings*, Dec. 2006.

[3] Keser, B. et. al., "The Redistributed Chip Package: A Breakthrough for Advanced Packaging," *2007 Electronic Components and Technology Conference*, pp. 286-291,

[4] J. Sabatini, "GE's High-Density Overlay Technology," *Surface-Mount Technology*. Vol. 6, No. 3. Mar. 1992, pp. 18-19.

[5] Daum, W. et. al., "Overlay High-Density Interconnect: A Chips-First Multichip Module Technology," *Computer*, vol. 26, no. 4, April 1993, pp. 23-29.

[6] Meyer, T. et. al., "eWLB System in Package – Possbilities and Requirements," *IWLPC Proceedings*, Oct. 2010, pp. 160-166.

[7] Kang, I.S, et. al., "Wafer Level Embedded System in Package (WL-ESiP) for 3D SiP Solution," *IWLPC Proceedings*, Oct. 2010, pp. 153-159.

[8] C. Scanlan et. al., "Adaptive Patterning for Panelized Packaging," *IWLPC Proceedings*, Nov. 2012.

[9] JEDEC Standard JESD22-A104C, Temperature Cycling, 2005.

[10] R. Darveaux et. al., "Reliability of Plastic Ball Grid Array Assembly," Ball Grid Array Technology, J. Lau Editor, McGraw-Hill, New York, 1995, pp. 379-442.

[11] R. Darveaux, "Effect of simulation methodology on solder joint crack growth correlation," *50th ECTC Conf. Proc.*, 2000, pp. 1048-1058.