REWORK AND REBALL CHALLENGES FOR WAFER-LEVEL PACKAGES

Lauren Cummings and Priyanka Dobriyal, Ph.D.
Intel Corporation
Hillsboro, OR, USA
lauren.cummings@intel.com

ABSTRACT
With increasing consumer demand for smart phones, wearable devices, and Internet of Things applications, there is a growing trend in package and printed circuit board (PCB) miniaturization. In particular, wafer-level packages (WLPs) have garnered recent popularity for their affordable cost, small footprint, and thin profile. Component suppliers must be prepared to support failure analysis (FA) for PCB-assembled WLPs, including fault isolation (FI), non-destructive screening, as well as destructive analysis techniques. If a board- or package-level failure is subtle or cannot be detected non-destructively, the WLP requires rework and reball before proceeding with further component-level testing and destructive FA. Due to their fragility and small form factor, the rework and reball process steps pose considerable risks for WLPs. The component lacks a package substrate and is easily damaged using traditional rework tooling and handling. On high-density boards and modules, there is also a risk for adjacent board-side passives or packages to be bumped and damaged during package removal from the PCB. The present work addresses the rework and reball challenges of a specific WLP case study, and suggests improvements for maintaining the true failure signature. Rework and reball recipes were successfully developed for a WLP, and optical microscopy (OM) and C-mode scanning acoustic microscopy (CSAM) were used to inspect for thermally or mechanically-induced artifacts. By implementing enhanced WLP rework and reball methods, the industry will be better poised to improve the quality and reliability of small form factor devices.

Key words: Rework, reball, WLP, WLCSP, failure analysis

INTRODUCTION

There is a growing consumer demand for "smart" wearable products, including watches, fitness bands, eyewear, and headphones. Barriers to the market still exist, as many consumers are concerned with style, battery life, functionality, and cost. The mature smart phone market faces similar challenges, with consumers desiring thinner phones with longer battery life and increased sensor functionality. High-density and stacked PCB designs, system in package (SiP) assemblies, and small form factor packages have all emerged as solutions for sleeker styles and improved functionality. In particular, WLPs are gaining popularity for their low cost, small footprint, and thin profile. The newest smart phone models contain an average of 5-7 WLPs, with many WLPs used as RF transceivers, power management units, audio amplifiers, and BlueTooth and GPS modules [1].

Unlike traditional packages, WLPs are packaged and bumped first, then diced. Passivation and dielectric layers are added to the die frontside, followed by metallic redistribution layers. A second dielectric layer is deposited, then the underbump metallization and solder balls are attached. Lastly, the packages are singulated from the wafer. Figure 1 illustrates the two different categories of WLPs: the traditional a) fan-in WLP, and the newer b) fan-out WLP. The dielectric is exposed on the edges and frontside of fan-in WLPs, while the silicon backside is often covered with a protection tape. As the name suggests, fan-out WLPs "fan out" interconnects from the smaller silicon die to the larger package dimensions. The fan-out design allows for ball pitch customization, higher I/O density, and easy integration with SiPs and other multi-die packages. The board-level reliability is also improved by protecting the silicon die with an epoxy mold compound.

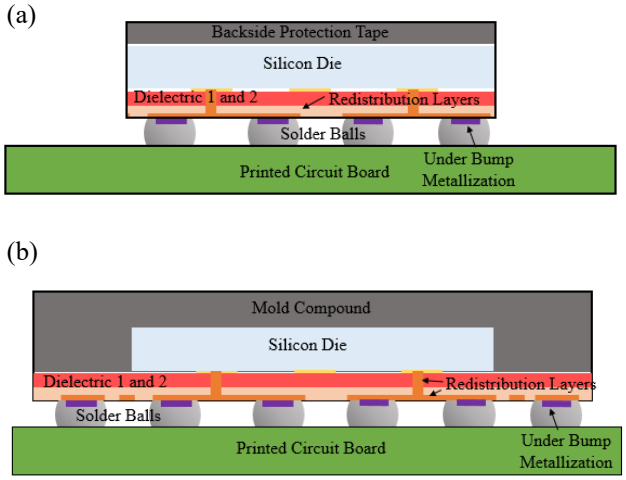


Figure 1. Cross-section schematics of a **a)** traditional fan-in wafer level package, and a **b)** newer fan-out wafer level package.

Though WLPs offer considerable advantages, they also pose challenges for failure analysis—particularly when reball and component-level testing are required. Figure 2 illustrates the typical FA process flow for a failing system, such as a mobile phone, tablet, or wearable device. Fault isolation is performed first to identify the failing component, then non-destructive FA is used to inspect for failures at both the board- and package-level. If the failure is subtle or cannot be found non-destructively, the package must be reworked

and reballed before proceeding with socketed component-level testing and further FA. Due to their small form factor and fragility, WLPs can prove particularly difficult for the rework and reball process steps. The present work provides an overview of the FA process flow, emphasizing component rework and reball methodologies. The challenges of a specific WLP case study are discussed, and rework and reball improvements are implemented in order to minimize thermally- or mechanically-induced artifacts.

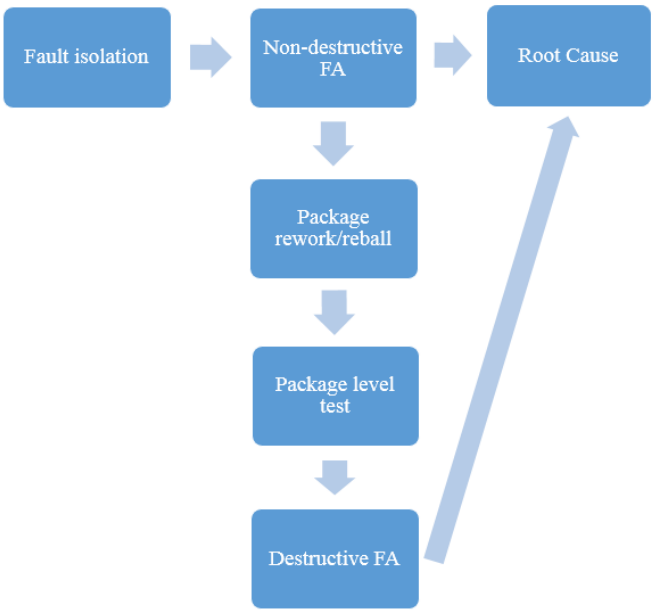


Figure 2. Diagram showing the typical process flow for system-level failure analysis.

Fault isolation is an important first step in the system-level FA process. By isolating the failure to a specific package or interconnect, an optimal FA approach can be assessed and throughput time can be greatly reduced. Several fault isolation techniques can be used to measure for opens and shorts, including hand probing, time domain reflectometry and the newer electro optic terahertz pulse reflectometry [2]. High-resistance shorts can also be detected by powering the board and measuring the localized temperature increase using infrared thermal imaging techniques [3].

After the failing component has been isolated, non-destructive FA is performed to inspect for gross board- or package-level failures. OM can be used to inspect for external board or package defects, such as foreign materials or superficial die and overmold cracks. Since WLPs have exposed dielectric layers (and since fan-in WLPs have exposed bulk Si), it is especially important to perform a first pass inspection for external chips or cracks. It is recommended that optical inspection be performed again after rework and reball, to confirm that no artifacts were introduced.

CSAM, a popular non-destructive technique, can be utilized to detect internal defects or to evaluate the extent of external damage. CSAM uses an ultrasound transducer to raster-scan the package backside. At material interfaces, an acoustic pulse is reflected back to the transducer and recorded as signal amplitude. Air-solid interfaces occur at the locations of cracks, voids, or delamination, and return high-intensity reflections. CSAM is thus a valuable metrology for identifying gross internal package defects that cannot be detected with simple optical inspection. CSAM can also be used to screen for rework and reball artifacts, particularly thermally-induced delamination.

2D X-Ray is another common technique that provides an effective "quick pass" inspection for board-level solder defects, such as voids and bridging. By optimizing the sample tilt and rotation angles, more subtle non-wet open and non-contact open defects can also be detected. However, 2D X-Ray is not capable of detecting sub-micron defects, such as board-level solder or via cracks.

Recently, more advanced imaging metrologies have emerged as powerful non-destructive FA techniques. In particular, 3D X-Ray computed tomography has proven effective at detecting both board-level and package-level sub-micron defects [2,4]. Multiple 2D X-Ray images are collected as the sample is rotated at fixed angle increments. The 2D X-Ray images are then superimposed to generate a three-dimensional volume. The superimposed image can be manipulated to display virtual "slices" of the sample, allowing for inspection of the solder joints, via barrels, and traces.

If board-level or gross package-level failures are not detected using non-destructive techniques, the package must be sent for socketed component-level testing and further FA. Standard test sockets use a spring-loaded floating base that is guided by the solder balls instead of the package edge. Reball is thus required in order to align the package and enable good electrical contact with each pin. It is very important to preserve the defect signature prior to testing; accordingly, precautions must be taken to improve the reball yield and reduce thermal and mechanical artifacts.

Rework and reball process yield is influenced by several different factors, and can be classified into four general categories, including 1) personnel, 2) methods, 3) materials, and 4) machine/tooling [5]. Figure 3 summarizes the various categories and sub-categories for a typical PCB assembly.

The personnel category represents the "human factor," and includes handling, training, and quality control. The personnel factor is especially important for WLPs and other small form factor devices, as the components are much more fragile than standard flip-chip packages. Proper handling must be used in order to minimize mechanical artifacts and prevent damage to the bulk Si or dielectric layers.

The methods category encompasses the process steps for rework and reball, including sample preparation, package removal from the board, solder removal from the package-side pads, and reball of the package. Prior to demount, the sample must be prepared by removing any heat spreaders, thermal grease, corner glue, or underfill. The package is then demounted from the PCB using either mechanical or thermal methods. Next, solder is removed from the package-side pads using solder wicking with a braided wire and solder tip, or using a no-contact vacuum scavenging technique [6]. Lastly, the package is reballed using either a stencil, preform, or laser jetting method. Figure 4 shows the transformation of a WLP throughout the rework and reball process.

Materials—such as package type, PCB design, flux, underfill, and corner glue— also influence the rework and reball process yield. With higher board densities and package miniaturization, it becomes increasingly difficult to selectively heat and remove parts from a small footprint. Since WLPs do not possess a package substrate or solder mask, there is also an increased risk to damage the dielectric or even the metal redistribution layers.

Lastly, rework yield is impacted by the machines and tooling used to handle and process the package. Rework machines can vary greatly in cost and complexity—ranging from a hot air pencil and tweezers, to a fully automated rework station. Semi-automated and fully automated rework tools are expensive, but can greatly minimize the risk for thermal and mechanical artifacts. The latter is especially important for WLPs, as it is difficult to handle and secure the parts.

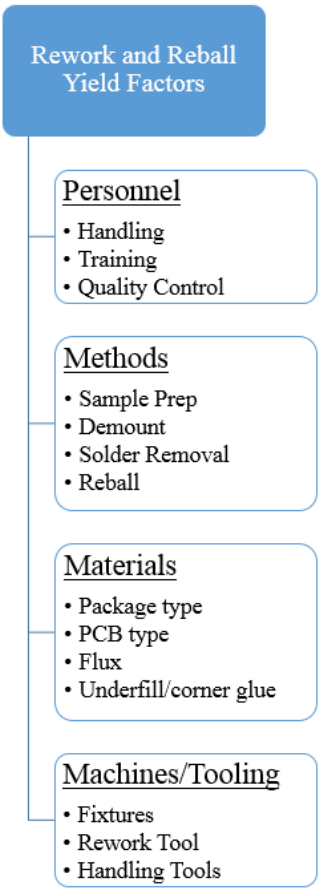


Figure 3. Diagram showing the factors that dictate the yield of rework and reball processes.

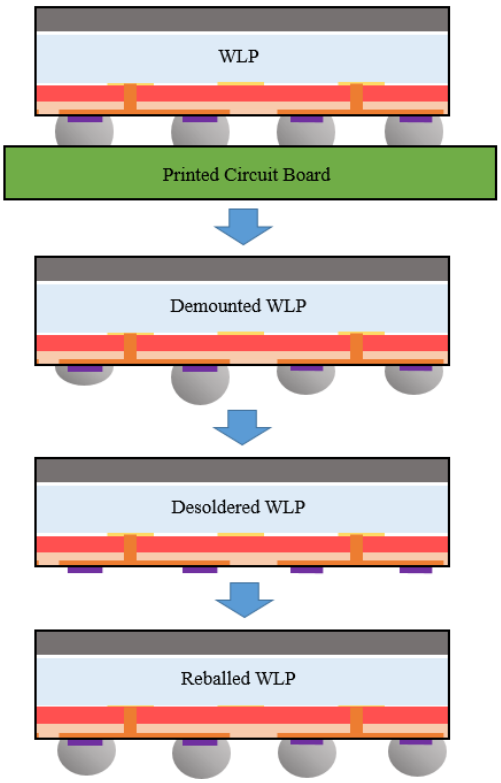


Figure 4. Schematic illustrating the WLP condition after demount, desolder, and reball.

EXPERIMENTAL

A 3.5 x 3.5 fan-in wafer-level chip scale package was used for rework and reball evaluation. The package uses standard 0.5 mm pitch SAC405 solder balls. Demount, desolder, and reball steps were all performed using a semi-automated rework system. OM and CSAM inspection were performed before and after to assess for mechanically- or thermally-induced rework artifacts.

Package Demount from Board:

Figure 5 shows the set-up for package demount using a semi-automated rework system. The PCB assembly was clamped on a motorized XY stage and a vision system was used to manually center the WLP between top and bottom heaters. Top heater convection was localized using a 10 mm nozzle attachment. The nozzle was lowered onto the board, and an automated recipe was used to heat the part to reflow temperatures. To prevent thermal artifacts and ensure complete solder liquidation, the WLP was heated between 217 and 250 °C for 60-90 s. At peak reflow of 235-250 °C, the WLP was lifted from the PCB using vacuum suction through a metal pick-up tube. Figure 6 shows a 3.5 mm vacuum cup that was used to minimize mechanical stress to the WLP silicon backside.

In order to accurately measure the WLP solder joint temperature, a board assembly was sacrificed for temperature profiling. A WLP was demounted using a test recipe, and holes were drilled through the WLP board-side pads at the center, top right, and bottom left corners. Thermocouples were inserted through the backside of the PCB, until flush with the board-side pads. The thermocouples were then bonded and cured using a thermally-conductive epoxy. After the thermocouples were attached and tested, the PCB assembly was clamped to the X-Y stage of the rework system. A thin coating of flux was applied to the WLP board-side pads, and a vision system was used to pick, align, and place a fresh WLP on to the PCB. The package was reflowed to the board, and heater times and temperatures were iteratively adjusted until the critical reflow parameters were satisfied.

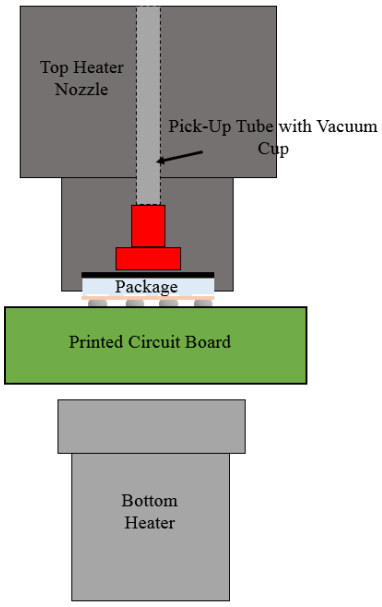


Figure 5. Schematic of set-up used to demount a WLP from a PCB.

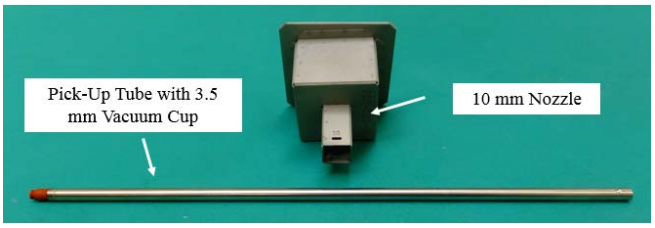


Figure 6. Photograph showing the 10 mm nozzle, metal pick-up tube, and 3.5 mm cup used during vacuum pick-up of the WLP.

Package Desoldering:

After package demount, residual solder debris must be removed in order to provide a smooth and even surface for reball. Conventional desoldering is performed using the solder wicking technique with a hot air pencil and braided wire. Alternatively, package desoldering was performed using a no-contact vacuum scavenging method that eliminates mechanical stresses and ensures repeatable temperature control. Figure 7 shows a schematic of the vacuum scavenge set-up, including the vacuum nozzle and custom fiberglass-resin fixture. The WLP was placed front-side up in a spring-loaded fixture and clamped to the rework stage. The top and bottom heaters were used to heat the component above solder liquidus temperatures while a

vacuum nozzle descended 0.1-0.25 mm above the component. The molten solder debris and flux were then vacuum-suctioned in a pre-programmed XY raster pattern. Because the WLP is so small, only two passes were required to remove the residual solder from the package-side. The scavenging height and velocity, scavenge pattern, and top and bottom heater temperature were all controlled using a semi-automated recipe. To reduce artifacts and ensure liquidus temperatures, a test WLP was profiled by attaching a thermocouple to the component backside.

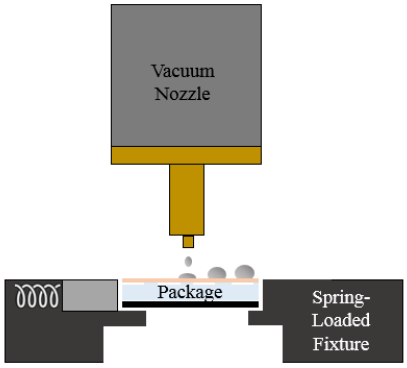


Figure 7. Schematic demonstrating the tool set-up for vacuum scavenging a WLP.

Package Reball:

Figure 8 shows a schematic of the preform process used to reball the WLPs. After desoldering the package-side, residual debris and flux were cleaned using a flux-off spray and a coarse-bristled brush. A custom solder preform was placed ball-side up on a fiberglass-resin fixture. A thin coat of flux was applied to the package-side, and the WLP was placed on top of the preform. To improve ball attach and prevent the part from blowing away, a metal weight was placed on top of the WLP and preform. The weight, WLP, and preform stack were reflowed using the rework system, then the weight and preform were carefully removed after the reballed package had cooled.

The entire stack-up was profiled by attaching a thermocouple between the preform and the fixture. The top and bottom heater temperatures and times were adjusted in order to meet the following critical process parameters: 60-90 s soak between 150 °C and 217 °C, and 60-90 s reflow between 217 °C and 250 °C.

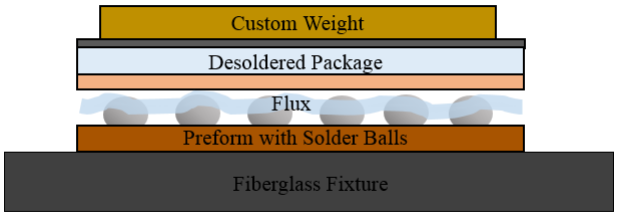


Figure 8. Schematic showing the reball stack-up.

RESULTS AND DISCUSSION

Several board-assembled WLP units were used for rework and reball development activities. A demount recipe was created by profiling a PCB-mounted WLP and adjusting the top heater and bottom heater settings. Figure 9 shows the temperature profile obtained during demount. WLPs were successfully demounted using a peak temperature of ~235 ° C and a 70 s time above liquidus.

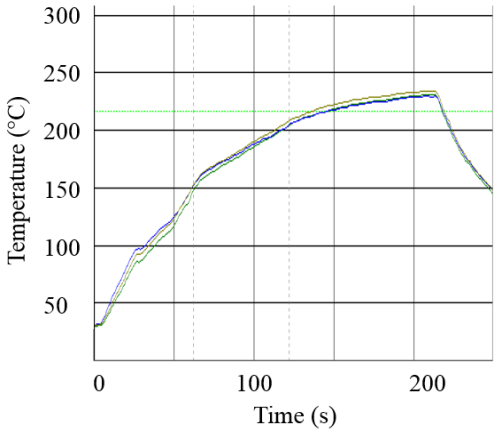


Figure 9. Graph showing the package temperature as a function of time during demount.

After demount, a recipe was developed to vacuum scavenge the solder from the package. A WLP was placed front-side up in a spring-loaded snugger inside a fiberglass-resin composite fixture. A thermocouple was attached to the silicon backside and used to measure the package temperature. Top and bottom heater set points were adjusted until the package temperature remained between liquidus and 240 °C. Figure 10 shows a representative microscope image of one of the desoldered WLPs. Inspection showed the recipe provided adequate solder removal, with only a thin layer of solder remaining on the under bump metallization.

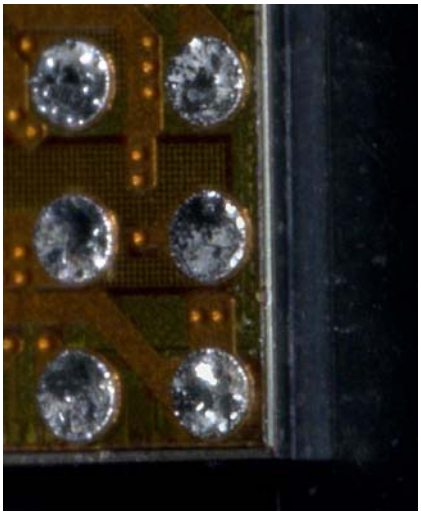


Figure 10. Representative microscope image showing WLP after successful vacuum scavenging. Only a thin layer of solder remains on the under bump metallization.

Lastly, a reball recipe was created. The preform-package-weight stack-up was placed on a fiberglass-resin fixture, and a thermocouple was mounted between the preform and the fixture. Top and bottom heater settings were adjusted until the package achieved the critical time and temperatures for soak and reflow. Figure 11 shows the WLP temperature profile during reball, with a ~240 °C peak reflow temperature. Due to the low mass of the weight and WLP, the reball stack-up blew away during several reball attempts. Consequently, the air flow had to be reduced at the beginning and end of the recipe. Optical inspection showed the balls properly wetted to the under bump metallization, as shown in Figure 12. No opens or shorts were visible on any of the inspected units.

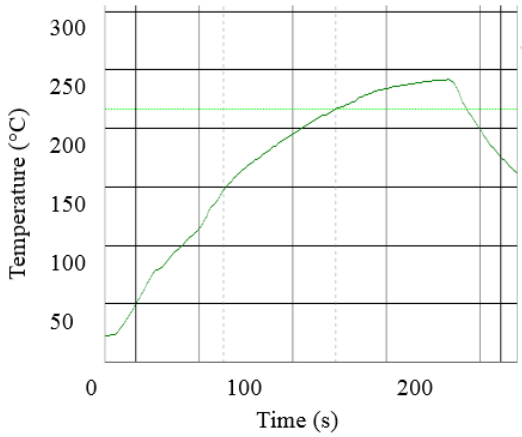


Figure 11. Graph showing the package temperature as a function of time during preform reball.

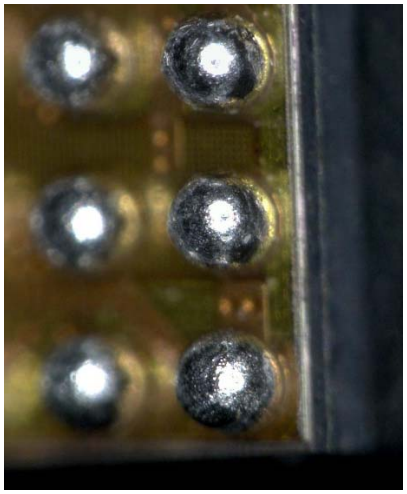


Figure 12. Representative microscope image showing WLP after successful reball.

CSAM and optical inspection were performed before and after the rework process. No thermal artifacts were found, but the following two mechanical artifacts were observed on multiple units: 1) peeling and chipping of the backside protective tape, and 2) chipping of the dielectric and top metal layers. Figure 13 a,b shows representative CSAM images of the package backside, captured before demount and after reball. Following reball, damage was detected at the interface between the backside protective tape and bulk silicon. Optical images confirmed that the backside protective tape chipped off and exposed the silicon, as shown in Figure 13c. Following a step-by-step investigation, it was determined that the backside protective tape was peeling and chipping after storing and removing the WLPs from adhesive packs. Though the tape chipping is cosmetic, care should always be taken to reduce artifacts and prevent masking of the true failure signature. Accordingly, the backside tape artifact was eliminated by storing subsequent samples in plastic trays instead of adhesive packs.

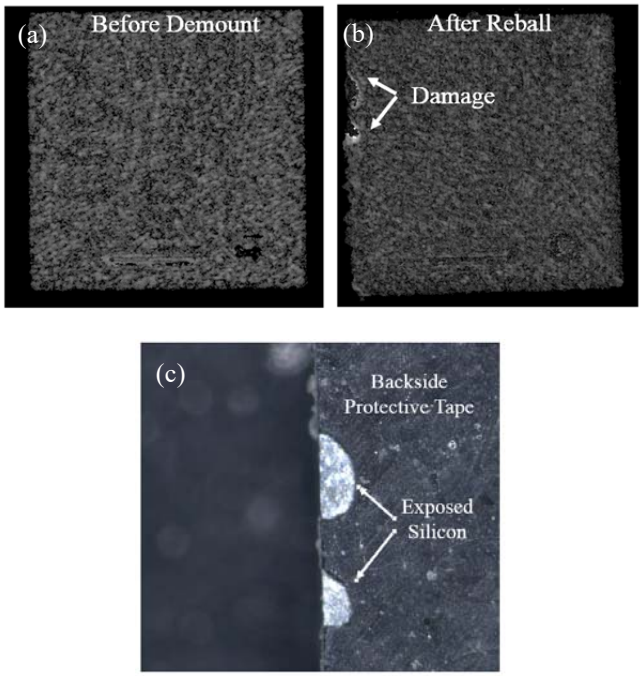
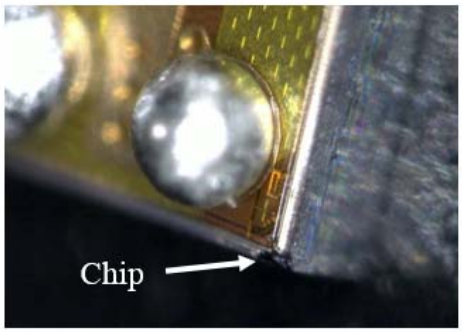


Figure 13. CSAM images showing the package backside a) before demount and b) after reball. An c) optical image shows chipping of the backside protective tape.

Frontside chipping was also detected during optical inspection. Figure 14 a,b reveals ~50 µm chips near the dielectric and top layer metallization. The chips were approximately the same diameter as the tips of fine point metal tweezers used during WLP handling. Chipped WLPs failed subsequent component-level test, suggesting that the artifacts affected the package integrity. Thus, the results show that frontside chipping can not only mask the true failure signature but also induce failures during component-level testing. All further chipping was eliminated by handling WLPs with a vacuum pen and plastic tweezers. Table 1 summarizes all the rework-related challenges and risks discussed in this paper, and lists the implemented solutions.

(a)



(b)

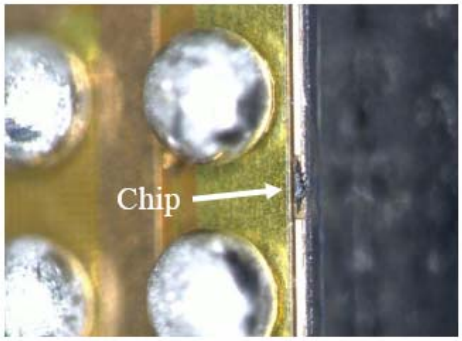


Figure 14. Representative optical images (a,b) showing chipping near the dielectric and top layer metallization.

Table 1: Summary of challenges and risks associated with each rework step, along with the implemented solution.

Rework Step	WLP Challenge/Risk	Implemented Solution
Demount	Damage to the package backside or edges; risk of bumping adjacent components on high-density PCBs	Thermally demount package using pick-up tube with a small, soft vacuum cup
Desolder	Thermal artifacts from overheating the package	Use temperature-controlled vacuum scavenge method
Reball	Mechanical damage when part blows away; low ball attach yield due to insufficient force for ball adhesion	Use low airflow during reball; use metal preform weight to increase mass of preform stack-up
Handling/Storage	Mechanical artifacts from metal tweezers and tacky packs	Handle packages with plastic tweezers or a vacuum pen; use plastic tray or low-adhesive pack to store and transport packages

CONCLUSION

Rework and reball recipes were successfully developed for board-assembled WLPs. OM and CSAM inspection were performed in order to evaluate the rework and reball process yield, and screen for thermally or mechanically-induced artifacts. The risk for mechanical damage was minimized by thermally demounting the WLPs with a vacuum pick-up tube and soft vacuum cup. Temperature was well-controlled using a no-contact vacuum scavenge technique to desolder the package. Lastly, the WLPs were reballed using a solder preform and a small metal weight. CSAM and OM did not reveal any thermal artifacts during the rework process, but chipping artifacts were found on the backside protection tape and near the dielectric and top layer metallization. WLPs with frontside chips failed subsequent component-level testing, showing that small dielectric and metal defects can sacrifice electrical functionality. Both backside and frontside chipping artifacts were successfully eliminated by improving handling and storage techniques. By implementing similar rework and reball improvements, the industry will be prepared to support WLP FA while maintaining the true defect signature.

REFERENCES

[1] TechSearch International, Inc., "The Future of Packaging and Assembly Technology," 2016.

[2] Y. Li, P.K.M Srinath, D. Goyal, "A Review of Failure Analysis Methods for Advanced 3D Microelectronic Packages," Journal of Electronic Materials, January 2016, 116-124.

[3] A. Wilson, "IR Cameras Tackle PCB Inspection Applications," 2007.

[4] M. Pacheco, "Detection and Characterization of Defects in Microelectronic Packages and Boards by Means of High-Resolution X-Ray Computed Tomography," IEEE Electronic Components and Technology Conference, June 2011, 1263-1268.

[5] A.A. Primavera, "Process Issues for Fine Pitch CSP Rework and Site Scavenging," EtroniX Conference Proceedings, 2001.

[6] J. Wade, R. Aspandiar, D. Naugler, T. Leahy, "PCB Pad Site Dress Methods on BGA and Socket Pad Arrays," SMTAi Conference, 2010.

ACKNOWLEDGEMENTS

The authors would like to thank the teams from the Global Platform Analysis Center (GPAC) and Assembly and Test Technology Development (ATTD) FA Labs.