A Novel Three Dimensional Wafer Level Chip Scale Packaging Technology-Manufacturing Process Development and Reliability Characterizations

by

Zhaozhi Li

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 9, 2010

Key Words: 3D Packaging, Die to Wafer Integration, Silicon on Silicon, Pb-free Reflow, Capillary Underfill, No Flow Underfill, Fine Pitch Flip Chip Assembly

Approved by

John L. Evans, Chair, Thomas Walter Professor of Industrial and Systems Engineering
Daniel F. Baldwin, President of Engent Inc., Norcross, GA
Richard F. Sesek, Assistant Professor of Industrial and Systems Engineering

Abstract

Three Dimensional (3D) packaging has moved to the forefront in the electronic packaging industry, as the trend toward higher performance, smaller form factor and lower cost continues.

The development of the manufacturing technology for a novel Three Dimensional Wafer Level Chip Scale Packaging (3D-WLCSP), that leverages the existing infrastructures of high throughput wafer level packaging and low cost flip chip assembly process, is conducted in this research.

Two levels of device packaging integration are involved in this novel 3D packaging architecture. The first-level integration is realized by the face-to-face bonding of a thin profile, fine-pitch, flip chip die to a wafer formatted Wafer Level CSP substrate. The second-level integration features the assembly of the first-level packaged component to an FR4 organic substrate (PCB). The 3D-WLCSP packaging technology developed in this research utilizes a 3D die to wafer integration methodology that provides a cost effective, rapid time to market 3D packaging solution.

Research efforts were focused on the high-density flip chip wafer level assembly techniques for the packaging of 3D-WLCSP, as well as the challenges, innovations and solutions associated with this type of packaging technology. In this work,

the flip chip pitch and bump size are varied as well as the key assembly materials (including fluxes and underfills) used to attach the flip chip to the WLCSP. Processing challenges and innovations addressed include flip chip fluxing methods for very fine-pitch and small bump sizes; impacts of reflow profile parameters and effects of reflow ambient gas environment toward silicon on silicon flip chip assembly yield, solder joint quality and reliability; wafer level flip chip assembly program setup and yield improvements; and the CSP solder joint voiding issues for the second-level assembly. Various aspects of the die-to-wafer assembly process are explored including scaling issues with high volume wafer level assembly, utilization of low cost underfill approaches such as no-flow underfills, and underfilling a solder balled WLCSP wafer with chip components in close proximity. Different reliability testing methods were utilized to evaluate the reliability performance of the packaged first-level and second-level assemblies.

This research has demonstrated that the 3D-WLCSP can be processed with high yield and can successfully undertake harsh environments with long-term, high reliability performance. The 3D-WLCSP is a qualified packaging architecture for the Pb-free 3D die-to-wafer integration.

Acknowledgments

First of all, I would like to express my deep appreciation and sincere thanks to my advisor Dr. John L. Evans, Thomas Walter Professor from Department of Industrial and Systems Engineering at Auburn University, and my co-advisor Dr. Daniel F. Baldwin, previously associate professor at Georgia Institute of Technology and currently president and CEO of Engent Inc., for their instruction, inspiration, guidance, encouragement and patience in the completion of this research and dissertation. I would also like to thank my committee member Dr. Richard Sesek from Department of Industrial and Systems Engineering at Auburn University, for his advice on my proposal and dissertation completion. In the meanwhile, I would like to express my thanks to Dr. Robert Dean for his proof reading of this dissertation and his suggestions.

The research project was sponsored by Engent Inc. The test vehicles were supplied by FlipChip International LLC. Therefore, I am grateful to both Engent and FlipChip International, for their great support in this research.

Most of this research project was conducted in the world-class facility at Engent Inc. I greatly appreciate the help from the talented engineers at Engent. Lehang La at Engent's FA lab trained my skills on the analytical equipment. Paul Houston, Brain Lewis and Keck Pathammavong at the advanced processing lab taught me how to use the SMT equipment and offered me tremendous help during my completion of this research

project. Nelson Brock helped me with the underfill jet dispenser. And the program managers as well as the technicians at Engent all extended their help to me without any hesitation. My friend Sangil Lee who graduated from Georgia Tech also offered me many helps during my research. Therefore, I am greatly indebted to them.

Finally and most importantly, I want to thank my parents, Guangxin Li and Zhenzhen Gao, my parents in-law, Chunsheng Yu and Airong He, and especially my wife, Jie Yu, for their encouragement, support and patience on my PhD work.

Table of Contents

List of Figures

List of Tables

CHAPTER 1 INTRODUCTION AND LITERATURE REVIEW

1.1 Trend of Electronic Packaging

The electronic packaging, which is also called semiconductor device assembly, or simply assembly, is mainly focused on the establishment of appropriate electrical interconnections and housing for electrical circuitry. Electronic packages have four key functions: electrical signal interconnection, mechanical protection, environmental protection and thermal management [1].

It is the ever-growing demand for “smaller, lighter, faster, higher complexity and density, lower power consumption, and cheaper technology” that drives the whole electronics industry [2]. The higher expectations from the customer for better and cheaper electronic devices, especially with mobile handsets, also call for product miniaturization with higher quality and lower cost. To meet these demands, the electronic packaging industry is keeping pace by developing new package solutions for these smaller, faster, lighter and lower priced electronics.

As the trend of electronic packaging goes towards smaller size, higher performance and lower cost, recent years have witnessed a shift from wire-bonded integrated circuits to flip chip interconnect structures. [1]

1.2 Flip Chip Packaging Technology

1.2.1 Introduction

The first flip chip package was introduced by IBM in the 1960s as the “Controlled Collapse Chip Connection” (C4) by putting a semiconductor device facing down toward ceramic carriers for the building of mainframe computers [3]. Today, flip chip packages are widely seen in watches, cell phones, GPS, hearing aids, and automotive engine controllers as well as computers.

Flip chip is actually not the name of certain packages but rather a package assembly processing method [4] or packaging interconnection technology. Flip chip describes the connection of an Integrated Circuit (IC) device to chip carrier such as Printed Circuit Board (PCB) with the active side of the chip facing toward the carrier. The active IC die is “flipped” during the assembly process toward the substrate and attached to the substrate in this fashion.

Flip chip offers lots of benefits over traditional packaging methods such as wire bonding process in terms of thermal performance, electrical performance, cost and size [1,5,6,7]. Unlike the wire-bonding technology that utilizes a point-to-point process to build chip-to-substrate connection, normally with gold wire loops at peripheral area of package, flip chip process is a Direct Chip Attach (DCA) method that forms all the interconnections between the die and substrate with the conductive bumps (usually solder balls). Moreover, the inductance of chip-to-substrate connection through flip chip bumps is much lower than that through wire loops [1]. Another advantage that flip chip holds over wire bonding technology is that flip chip packages provide higher packaging density

(more I/O in the same area), as flip chip has area array bump connections while the wire bonding package normally has only peripheral interconnections. As a result, flip chip can achieve higher signal density with reduced die and package size, and thus with a lower silicon cost. Some of the advantages and disadvantages of flip chip technology are listed in Table 1-1.

Table 1-1 Advantages and Disadvantages of Flip Chip Technology

Advantages	Disadvantages
✓ High packaging density, flexible I/O, area array structure, fine-pitch and small form factor ✓ High electrical and signal transmission performance ✓ Forming all interconnections in one solder reflow step to achieve low cost and high throughput	✓ Dedicated oven for cure of underfill to overcome the CTE mismatch between Die and Substrate ✓ Difficult to rework if underfilled ✓ Strict assembly process control ✓ High requirement for flatness of substrate

A schematic view of flip chip assembly is shown in Figure 1-1. A typical flip chip packaged component is composed of several functional parts – the die (chip) and substrate (chip carrier), Under Bump Metallization (UBM), flip chip bumps, substrate metallization (pads), bond metallurgy between pad and bump (normally Intermetallic Compound-IMC) and underfill encapsulant.

Silicon is the mostly commonly used material to make flip chip dies (IC). Flip chip die can also be made of gallium-arsenide (GaAs), indium-phosphide (IP), silicon-germanium and other materials [1,6].

There are three general stages for the flip chip packaging process. The first stage is the bumping of the flip chip die [5]. The second stage is the attachment of the bumped flip chip to the substrate to form the solder joint interconnects. The third stage is filling the gap between chip and substrate with electrically non-conductive epoxy material (underfill).

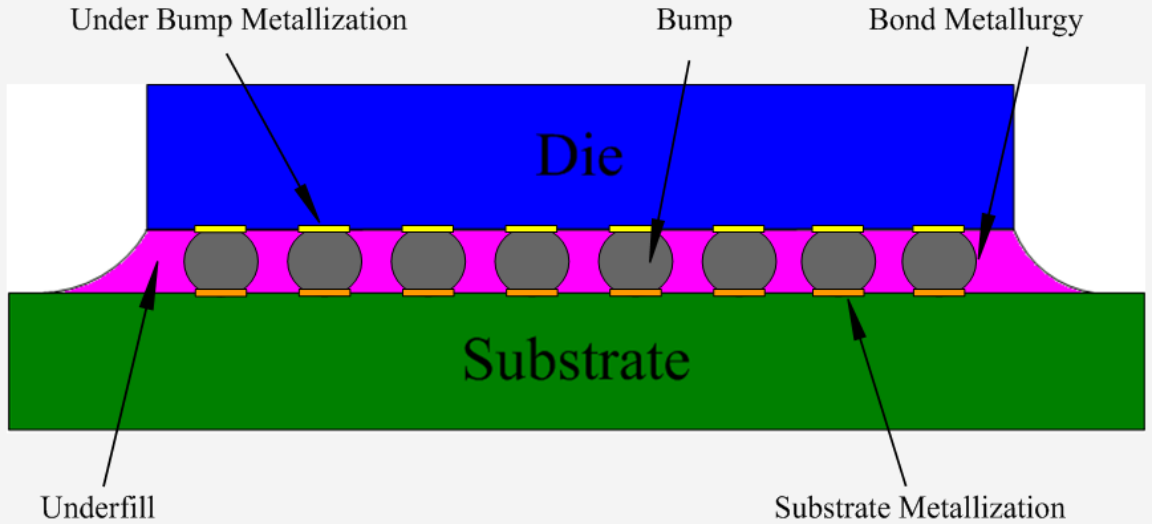


Figure 1-1 Schematic View of Flip Chip Assembly Configuration

The flip chip bump is connected with the final chip metallization (mostly aluminum) through the under bump metallization (UBM). A schematic view of flip chip UBM is shown in Figure 1-2. The chip passivation (mostly silicon nitride) is used to protect the chip from oxidation, moisture and corrosion [1]. A polyimide or benzocyclobutene (BCB) layer is deposited after the passivation layer for stress relief

purposes. The adhesion layer, usually made of chromium (Cr), titanium (Ti), tungsten (W) or Nickel (Ni), helps to hold the chip passivation, chip metallization, and stress relief layers tightly together. A barrier layer usually made of chromium, tungsten, chromium-copper, or nickel is deposited next to the adhesion layer to prevent diffusion of metal species and ionic contaminants from the wetting layer or other outside layers into the base chip metallization and the adhesion layer [1]. A wetting layer is placed next to the barrier layer to provide with the flip chip bump a consumable metal layer on which to wet and form the intermetallic compound (IMC). Copper (Cu) is the most commonly used material for the wetting layer. Other typical wetting layer materials include nickel (Ni), palladium (Pd) and platinum (Pt). The final UBM layer is the optional oxidation barrier, which is normally a very thin layer of gold. [1] The UBM layers are normally deposited by either evaporation or sputtering process.

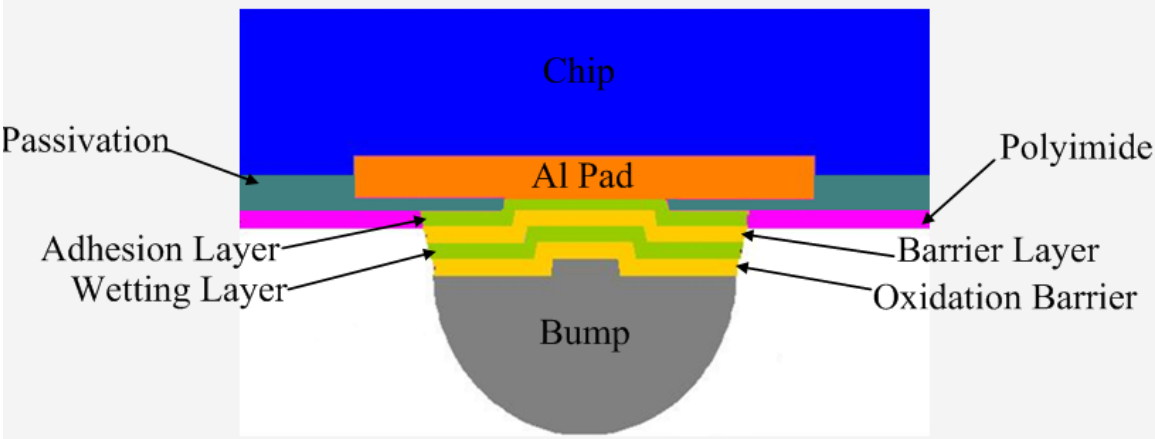


Figure 1-2 Structure of Flip Chip UBM

The flip chip bumps act as the electrical and signal transmission path between the die and substrate. The bumps also provide the thermal conduction path through which the

heat generated from the flip chip can be effectively dissipated to the substrate. In addition, flip chip bumps provide the mechanical support and environmental protection to the die. [1]

The bumps of flip chip can be categorized into several groups in terms of the bump materials and bumping process. These groups include solder bumps, wire stud bumps, and plated stud bumps [1]. Solder bumps are the most commonly seen flip chip bump formation, especially in low cost flip chip packages.

Several different methods can be utilized to deposit the solder bumps onto the flip chip UBM. Evaporation was the earliest method developed by IBM for the C4 technology. The evaporation method requires a metal mask that can be mechanically clamped and aligned to the wafer I/O pads. The solder is deposited onto the UBM by evaporating the high lead solder material and then reflowed to form the spherical bumps. The mask opening and registration determine the solder bump shape and pitch.

Electroplating process is a more often seen flip chip bumping method. Compared with the evaporation method, the electroplating process is less costly and more flexible, and is easier to process, especially for the high I/O count devices. The solder is electroplated through a polymer photopatterning mask to plate the solder material over the UBM. The volume and registration are controlled by the mask. The bumps are then reflowed to form the sphere.

A third method to deposit the solder bumps is through the solder squeegee printing process with stencil and screen printer. The bumps are reflowed after printing to form the sphere. At a pitch of 250um or lower, this process gets harder and more inaccurate due to the lack of high precision stencil down to that pitch [1].

Due to the fine-pitch characteristics of flip chip die, the fine-pitch, high density substrate is required to accommodate the chip. The more layers the substrate has, the more flexible and easier the substrate routing can be performed to achieve the required fine-pitch. In the meantime, the number of added layers will also induce the substrate fabrication difficulty and thus add to the fabrication cost.

The flip chip substrate can be made of ceramic, epoxy-glass laminate, polymer thin film buildup, resin-coated copper (RCC) buildup, glass, silicon, dielectric coated metal, liquid crystal polymer, metal matrix composite, low-temperature cofired ceramic (LTCC), ceramic thick film, multilayer high-temperature cofired ceramic (HTCC), paper, polyester, polyimide and so on. [1]

The substrate metallization (pad) is required to provide the flip chip bump with a flat surface to wet on and form the solder joint. For flip chip IC with solder bumps, the most commonly used pad metallizations include copper coated with an organic solderability preservative (OSP), copper plated with electroless plated nickel, and electroless nickel immersion gold (ENIG). Electroplated nickel and electroplated gold are also seen as the substrate pad surface finish metallization for flip chip assemblies.

The interconnect between the flip chip bumps and substrate metallization can mainly be achieved through two different ways, solder interconnect or adhesive interconnect. [1]. Figure 1-3-a and Figure 1-3-b show the solder joint interconnect. Figure 13-a shows the process in which a high-lead solder bump is placed onto a layer of tin-lead (normally eutectic, 63% tin, 37% lead) solder paste that was printed on the substrate metallization. This process ensures certain solder joint stand off height between chip and substrate since the high lead solder bump has a much higher melting

temperature (usually around 260°C - 300°C) and thus does not collapse during the eutectic solder reflow that usually has a peak temperature of 225°C. Therefore, the flip chip standoff above substrate is maintained which allows for proportionately smaller pitches than eutectic (63Sn-37Pb) or Pb-free solder joints as shown in Figure 1-3-b. In comparison, the eutectic solder would collapse during reflow and create possible bridge between the solder joints if the bumps are too close together due to the fine-pitch. The other advantage a high lead solder interconnect has over the eutectic or Pb-free solder joint is the higher thermal fatigue life (better reliability) since the taller standoff height provides a lower CTE gradient between chip and substrate.

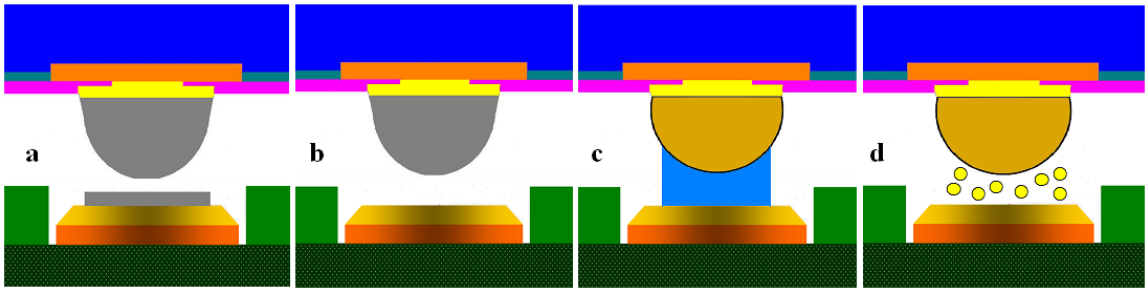


Figure 1-3 Flip Chip Interconnetion Methods

Figure 1-3-b shows the solder bump that have non-controlled collapse on the pad during reflow to form the solder joint interconnects. In this case, the solder bumps are heated to the temperature that is above the solder melting point. The height of solder joint after solder bump collapse depends on both the surface tension and volume of the liquidus solder as well as the pad sizes on both chip and substrate.

Figure 1-3-c shows the gold stud bump that is to be bonded on the substrate through isotropic conductive adhesives. Isotropic conductive adhesive materials are

typically made of thermosetting polymers (epoxy) filled with conductive particles such as silver flake particles. The adhesive material is fully conductive in all direction once cured and the gold stud bump is connected to the substrate pad through conductive particles. The isotropic adhesive can be applied by either printing the material onto the pad, or dipping the bumps into the thin layer of adhesive material.

Figure 1-3-d shows the gold stud bump interconnect by anisotropic adhesive material, which is normally seen in a film format (anisotropic conductive film-ACF) or a paste format. The film or paste contains conductive filler spheres such as metal (nickel or gold coated nickel) or metal-coated polymer (gold coated or nickel-gold coated polymer). In order to have these spheres be conductive, both certain levels of heat and compression force are applied through normally a thermal-compression bonder. The conductive spheres are compressed and captured between the bump and substrate pad to form the interconnects.

The original C4 packages used ceramic or other substrate with low coefficient of thermal expansion (CTE) due to the reliability issue caused by CTE mismatch between the silicon die and substrate [3]. Because the silicon die, Under Bump Metallization (UBM), solder bumps, pad and substrates have different thermal expansion coefficients, high strain and stress will occur on solder joints during thermal cycling or power cycling.

1.2.2 Flip Chip Underfill Introduction and Selection

For traditional flip-chip-on-board (FCOB) applications, the coefficient of thermal expansion (CTE) of the silicon device is about 3 ppm/°C while the FR4 organic substrate has a CTE of around 18 ppm/°C [8]. As shown in Figure 1-4, during thermal cycling or

power cycling, the solder joints are subject to stress and strain induced by CTE mismatch between IC and substrate. This will cause the solder joint fatigue which results in early solder joint failures. For flip chip on ceramic substrate applications, the CTE of ceramic substrate is around 6 ppm/°C. The CTE mismatch between die and substrate in this case is significantly lower than the flip chip on organic substrate application.

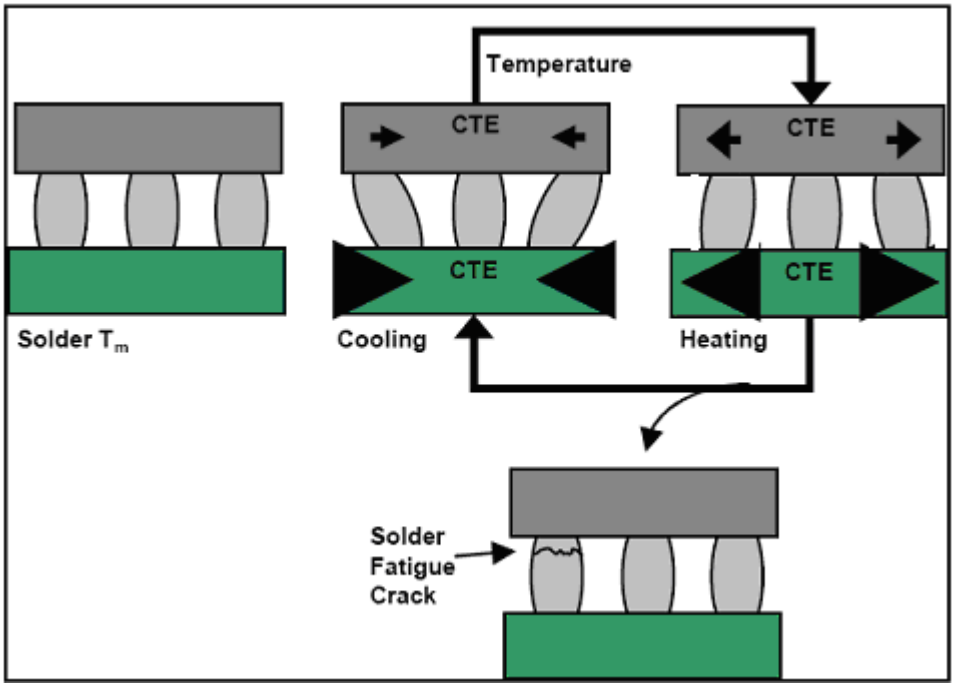


Figure 1-4 Flip Chip on Board Solder Joint Stress during Thermal Cycling [9]

In most applications of flip chip packages, underfill material, a specially engineered epoxy filled with silica particles, is applied to the gap between die and substrate to provide thermal- mechanical protection as well as environmental protection to the assembly [1,6]. The silica particles are filled to reduce the CTE of underfill. The underfill material, once cured, can wrap around and hold the solder joint, thus absorbing the stress and strain applied on the solder joints [7,10,11]. As a result, the thermal cycling

life of the solder joint and the finished package can be greatly improved, to even more than 10 times [12].

The effectiveness of underfill applied to flip chip package is mainly related to the relative difference in CTE between the chip and substrate, and also related to how well the underfill material wraps around the solder joint and maintains such a status during thermal cycling or mechanical impact situations. Therefore, the selection of underfill shall be based on the following factors – underfill glass transition temperature (Tg), underfill CTE both below and above Tg, viscosity and filler settling, flow characteristics, adhesion characteristics, compatibility with the flux, void formation, and so on. [13]

Ideally, the selected underfill shall have a CTE value that is reasonably close to the CTE of the flip chip solder joints, which is around 25 ppm/°C [14] for eutectic solder (63Sn/37Pb). An underfill CTE of 22-27 ppm/°C for flip chip package is recommended by Wong [15]. The ideal underfill candidate shall have the glass transition temperature above the package operation temperature or thermal cycling temperature, as the underfill CTE will normally increase 3-4 times once above Tg [14]. A short cure time, good flow characteristic, low moisture absorption, and good adhesion characteristic under a certain level of temperature and humidity are also desired for an ideal underfill candidate.

1.2.3 Flip Chip Capillary Flow Underfill Assembly Process

A conventional capillary flow flip chip process is shown in Figure 1-5. First, the flip chip die is picked up and dipped into a flux reservoir for a short time (usually within one second). The flux can also be applied onto the substrate pads directly by a screen printing process. The flux is used to clean the surface of solder bumps and pads, by

chemically breaking down the oxides. The flux can further prevent the formation of other oxidation during the forming of solder joints and support the solder wetting [1,5,6].

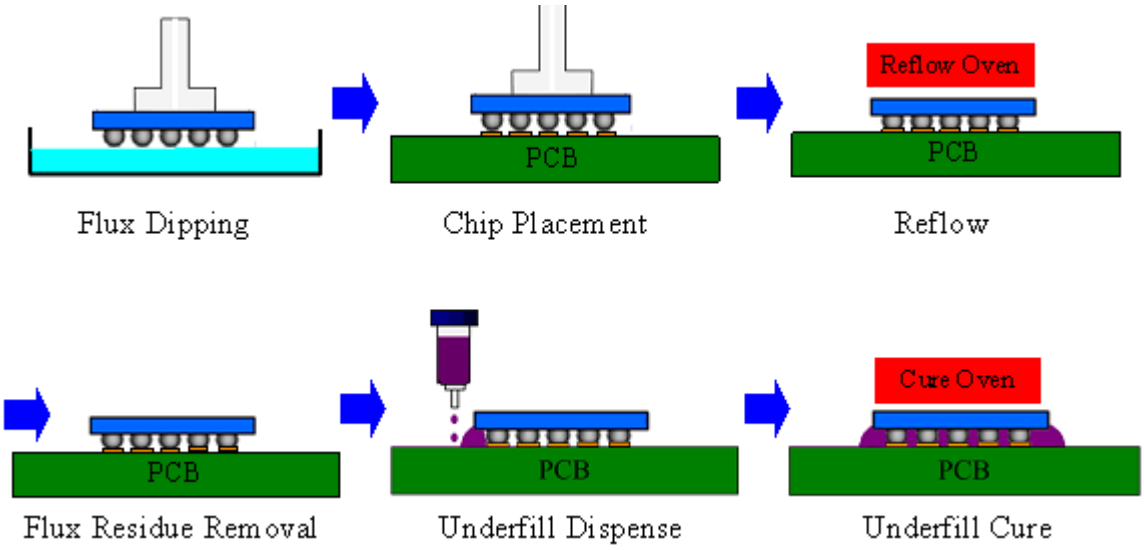


Figure 1-5 Flip Chip Assembly Process with Capillary Flow Underfill

After flux application, the chip is aligned with the substrate and placed onto the substrate by the high-precision placement machine so that each solder bump contacts its corresponding pad on the substrate. The package is then reflowed in a reflow oven following certain heating profile to make the solder wet on the substrate pad and form good solder joints that connect die and substrate, both electrically and mechanically.

After reflow, the flux residue is cleaned off as it can cause chemical corrosion on the solder joints. The flux residue may also cause the lack of adhesion and subsequent delamination of the underfill, which may undermine the package's long-term reliability performance. In some applications, the no-clean flux is utilized as the residue left by no-clean flux is inert and does not cause corrosion to the solder joints. However, the excessive residue left by no-clean flux may still cause the adhesion problem and

eventually delamination of the underfill. Therefore, flux/underfill compatibility evaluation through environmental testing (such as liquid-to-liquid thermal shock test) is normally performed to qualify the combination of no-clean flux and underfill before they are used for mass production. After the package reflow and flux residue removal, underfill is dispensed along the edge of the chip. Driven by capillary action, the underfill fills up the narrow gap between the chip and the substrate. In the final step the underfill is cured in a thermal batch oven. One critical factor for underfill to flow is the surface tension of the underfill relative to the surface energy of the die and substrate. When the surface tension of the underfill fluid is lower than the surface energy of the solid surface to be wet, the underfill will wet and flow. The lower the surface tension of the underfill, the lower the liquid contact angle with the solid surface and faster the wetting can take place. [14]

1.2.4 Flip Chip No-Flow Underfill Assembly Process

One drawback of the conventional capillary underfill flip chip process is its low throughput caused by time consuming process steps such as underfill capillary flowing, underfill curing, and possible cleaning of the flux residue. In addition, capital funding is required for the dedicated underfill cure oven. In comparison, the throughput of no-flow underfill process is higher due to its better match with the conventional SMT process and the elimination of many time-consuming steps such as flux cleaning, capillary flowing and cure of underfill [16-20].

Figure 1-6 shows the no-flow underfill flip chip process. First, a controlled mass of no-flow underfill is dispensed onto the substrate by means such as stencil printing,

jetting, or needle dispensing. A flip chip is then placed onto the substrate with a high precision pick and place system. The placement force acts as a compression on the no-flow underfill to allow the underfill to spread out and fill up the gap between chip and substrate as well as forming the fillets. In the final step, the assembly is reflowed through the reflow oven where the solder joint formation and underfill cure take place.

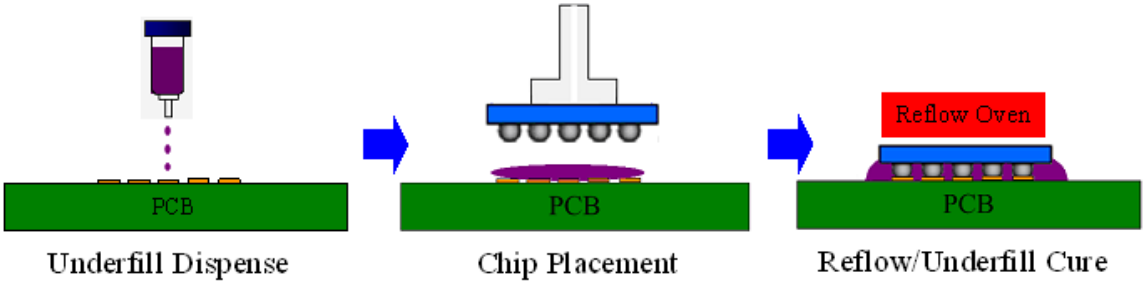


Figure 1-6 Flip Chip Assembly Process with No-Flow Underfill

Brian and Baldwin [20] estabilished a low cost flip chip technology based on selected no-flow underfill and showed that assemblies with no-flow underfill could survive 1000+ cycles of liquid-to-liquid thermal shock testing and have a mean time to failure exceeding 2000 cycles, thus proving that no-flow underfill is a reliable method for flip chip package assembly.

Sangil and Baldwin [21] conducted research on the fine-pitch flip chip assembly process with no-flow underfill focusing on the underfill void formation during reflow. It was found that a longer soak time can help reduce bubble growth.

1.3 Chip Scale Package (CSP) Technology

As the electronic packaging industry keeps moving toward smaller form factor as well as lower cost, Chip Scale Package (CSP) is gaining market share. A Chip Scale Package (CSP) is defined as an electronic package that has a total package size of less than 1.2 times the size of the chip [1]. CSPs offer a smaller footprint with reduced parasitics (especially at high frequencies) and provide greater inputs/outputs (I/O's) per unit area than traditional packages such as ball grid array (BGA) or thin small outline package (TSOP) [23]. The CSP package is often seen in SRAMS, DRAMS, flash memories, ASIC and microprocessors. Based on their structures, CSPs are generally categorized into 4 major types, as shown in figure 1-7 [18].

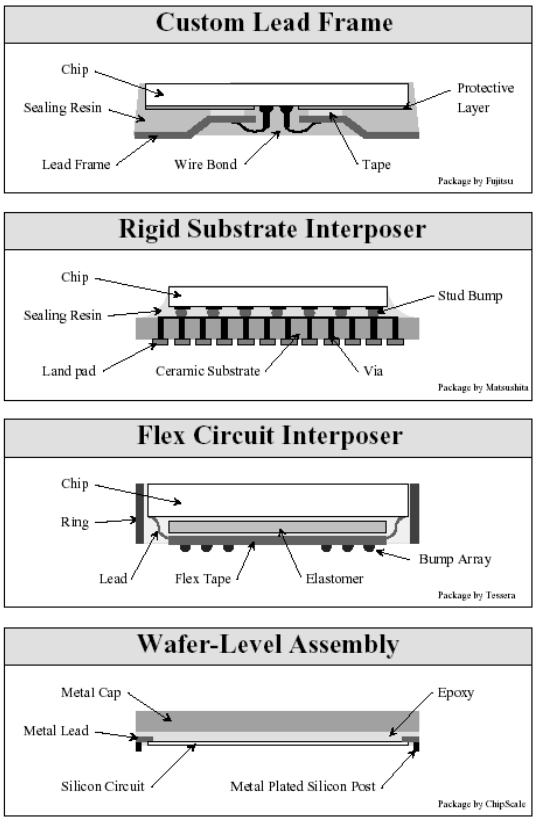


Figure 1-7 Main CSP Types [18]

1.4 Wafer Level Packaging

Wafer-level packaging (WLP) is defined as completing the packaging of a component at the wafer-level, as opposed to packaging at the discrete die level [1]. The WLP is a true CSP technology since the finished package has actually the same size as the die. With no additional processing needed at the die level, wafer-level packaging paves the way for true integration of wafer fab, packaging, test, and burn-in at wafer level for the ultimate streamlining of the manufacturing process from silicon to customer shipment [24]. Figure 1-8 shows the advantages of the Wafer Level Packaging [25].

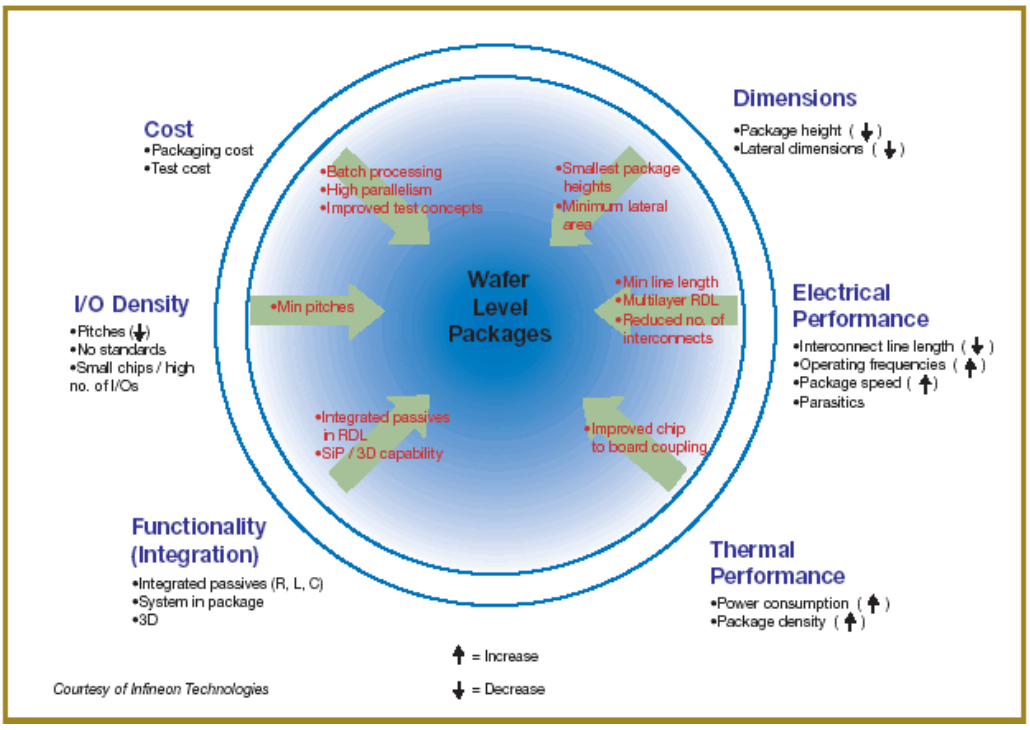


Figure 1-8 Advantages of Wafer Level Packaging

Although each step of WLP tends to be costly, the WLP process fabricates all die on the wafer simultaneously, thus resulting in an actual lower cost per package. WLP also

offers lower cost for the electrical test and burn-in test since those tests are done more efficiently at the wafer level. [25]

Philip Garrou [26] introduces several types of WLPs that produce fully packaged and tested chips on the wafer before dicing. The paper points out that for a given chip size, the non-underfilled WLP with larger solder balls (0.3-.05mm) tends to have better solder joint reliability than does the non-underfilled direct chip attachment with smaller solder balls (0.1-0.15mm).

John H. Lau [27] analyzed the cost issues associated with the WLP including IC chip yield, wafer-level redistribution yield, wafer-bumping yield, wafer-level underfill yield and so on. The paper also provided some useful equations to quantify the issues mentioned in terms of cost.

There has been a huge and steady growth in the market of WLP based on the advantages discussed before. Figure 1-9 shows the market demand of WLP that illustrates clear trend of growth [25].

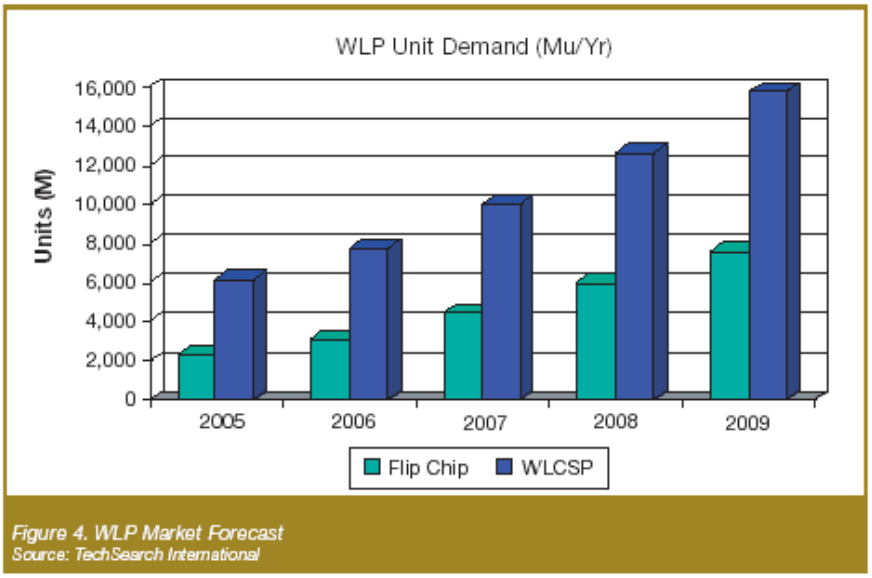


Figure 1-9 Market Growth of WLP

1.5 3D Packaging

With the increasing demand from the market for an even higher packaging density with multi-functional features, the three dimensional (3D) packaging came to the forefront. The 3D packaging is the packaging method that utilizes the Z dimension by stacking the die or packages vertically. It is a novel solution to package ICs like CPU, memory, sensors, logic circuits, etc. into a much smaller form factor than traditional 2D horizontal packaging, thus meeting the IC packaging trend of smaller, lighter, faster, higher complexity and greater density. Figure 1-10 shows the prediction of the 3D packaging technology in 5 years.

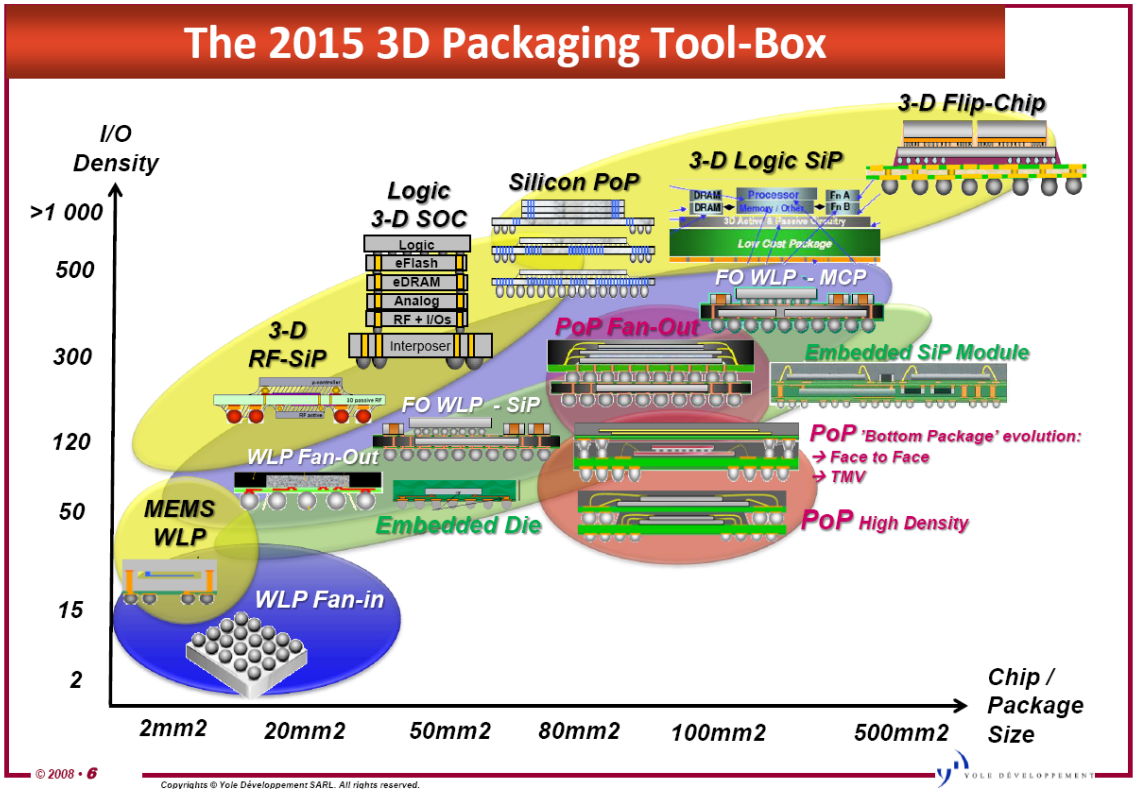


Figure 1-10 3D Packaging Tool-Box Predicted in 2015 [28]

Most of the current 3D packaging solutions can be grouped into four categories, in terms of the packaging structure and interconnect method. The four categories are: 3D die-stacked (die-to-die) packaging, 3D package-stacked (package-to-package) packaging, 3D wafer-to-wafer packaging and 3D die-to-wafer Packaging.

Figure 1-11 shows an example of 3D die-stacked package [29]. Two CSP dies were stacked together and connected to a chip carrier with wire bonding method. The stacked CSP technology enables the stacking of a wide variety of semiconductor components such as memory devices or flash drives to deliver the high level of silicon integration and small form factor required in portable multi-media products.

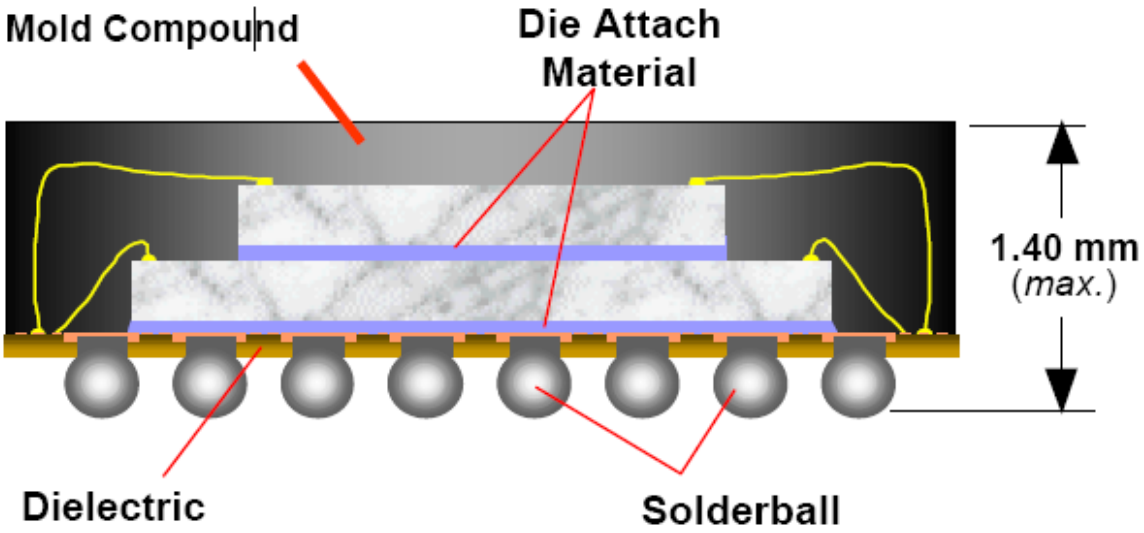


Figure 1-11 3D Stacked CSP Package [29]

A 3-D Package-in-Package (PiP) with 0.5mm pitch was introduced by Kamezos [30] as a 3D package-stacked packaging structure. In figure 1-12, three memory chips were stacked in an inverted LGA package which is on top of an ASIC, in a BGA

packaging formation. The packages were interconnected with wire bonding and then molded into a standard CSP as a whole final package.

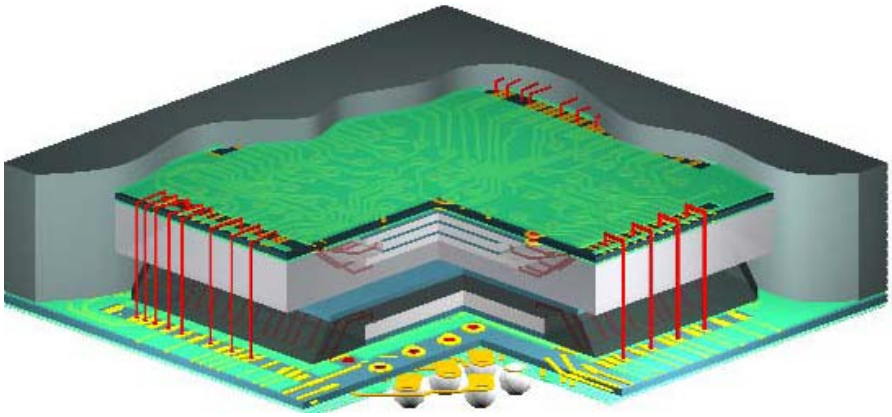


Figure 1-12 3D Package-in-Package Construction [30]

As described before, both the 3D die-stacked package and 3D package-stacked package achieve interconnections by utilizing the wire-bonding method. The wire-bonding technology, as discussed in chapter 1-2, utilizes normally the gold wire to build the point-to-point interconnection. The manufacturing throughput can be limited by such a process.

The 3D die-to-wafer and 3D wafer-to-wafer packaging technologies are emerging for the purpose of high manufacturing throughput as well as heterogeneous system integrations. Figure 1-13 shows a 3D die-to-wafer bonding method by EV Group [31]. The dies are placed on to the landing wafer and the die-to-wafer interconnections are realized by Through Silicon Vias (TSV), with the pressure and heat applied by a wafer-level bonding tool.

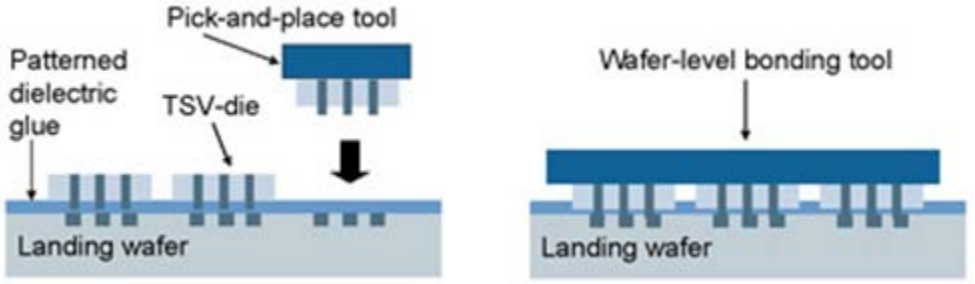


Figure 1-13 3D Die-to-Wafer Packaging [31]

Figure 1-14 shows the 3D wafer-to-wafer packaging structure. The wafers that have different function components fabricated are carefully aligned and bonded together with monolithic wafer-level BEOL-compatible process [51]. The vertical inter-wafer interconnection are achieved by through silicon vias. The bonded wafers are then diced and singulated so that the high manufacturing throughput of 3D chip stack packages can be obtained.

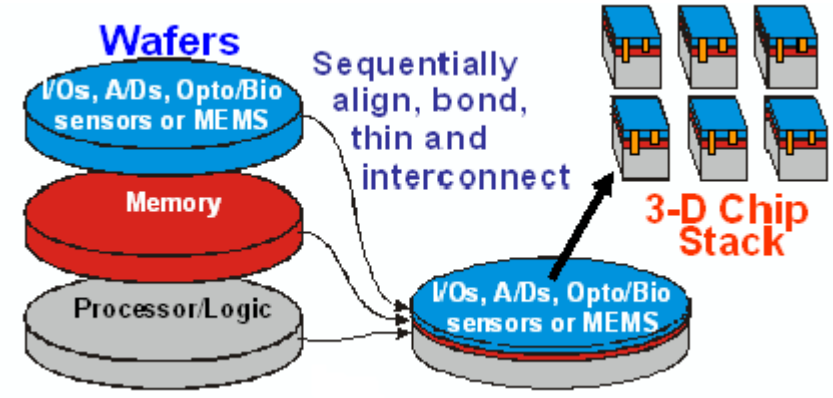


Figure 1-14 3D Wafer-to-Wafer Packaging [51]

Compared with the 3D die-stacked and 3D package-stacked packaging technologies, the 3D die-to-wafer and 3D wafer-to-wafer packaging technologies provide higher throughput, as the process is carried out at a wafer scale. However, both die-to-wafer and wafer-to-wafer packaging technologies discussed above utilize the through silicon vias technology. The TSV technology, though very efficient as the means of interconnect, is still in the research stage that is not a mature technology for mass production.

1.6 Introduction of 3D-WLCSP Packaging Technology

The manufacturing technology of a novel Three Dimensional Wafer Level Chip Scale Packaging (3D-WLCSP) is introduced in this research. Compared with the 3D die-to-wafer packaging discussed previously that utilizes the through silicon vias as interconnect method, the 3D-WLCSP realizes the die-to-wafer integration by leveraging both low cost flip chip process and the existing infrastructures of high throughput wafer-level packaging techniques. The active side of flip chip dies are face-to-face bonded to the active side of CSP components that are fabricated on the wafer, as shown in Figure 1-15. The die-to-wafer interconnects are realized by the fine pitch flip chip solder joints. High levels of packaging density, small package footprints, and thin package profile can be achieved through this packaging technology.

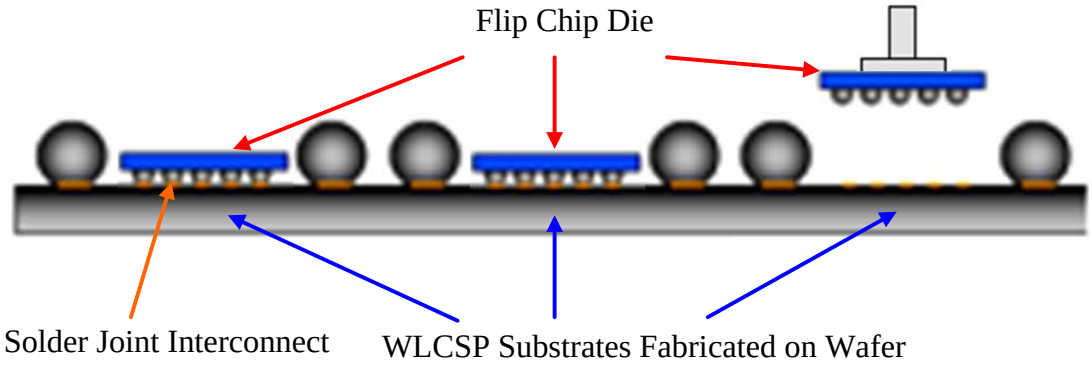


Figure 1-15 Flip Chip 3D Die-to-Wafer Packaging with 3D-WLCSP Technology

There are two levels of interconnections in the 3D structure. The first-level assembly features the thin profile, fine-pitch flip chip die to WLCSP integration on a wafer format, as shown in Figure 1-16. The flip chip and WLCSP are face-to-face connected through the flip chip solder joints. The solder balls fabricated at the peripheral area of the WLCSP are intended for the second level assembly.

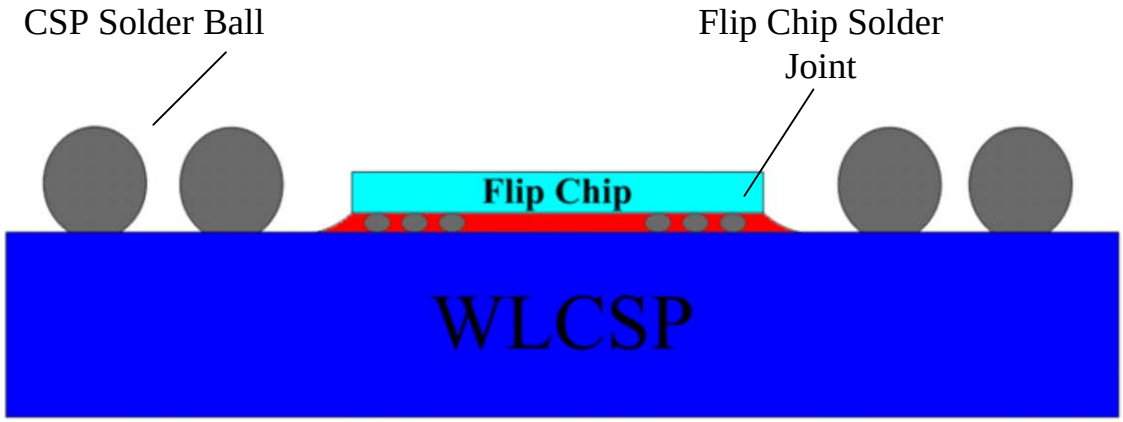


Figure 1-16 Schematic of a 3D-WLCSP Package First Level Interconnection

Figure 1-17 shows the conceptual view of the second-level interconnection for the 3D-WLCSP packaging structure. The second-level interconnection of the 3D-WLCSP architecture is achieved by surface mounting the first level package to the FR4 organic substrate (PCB) or other package carrier (silicon, ceramic, etc.). This level of assembly enables both the flip chip and CSP in the first-level package to communicate with other devices on the second level carrier. The second-level assembly is also compatible with the conventional SMT process and a standard CSP assembly process is adopted. Underfill can be applied at both first-level and second-level assembly to enhance the reliability performance.

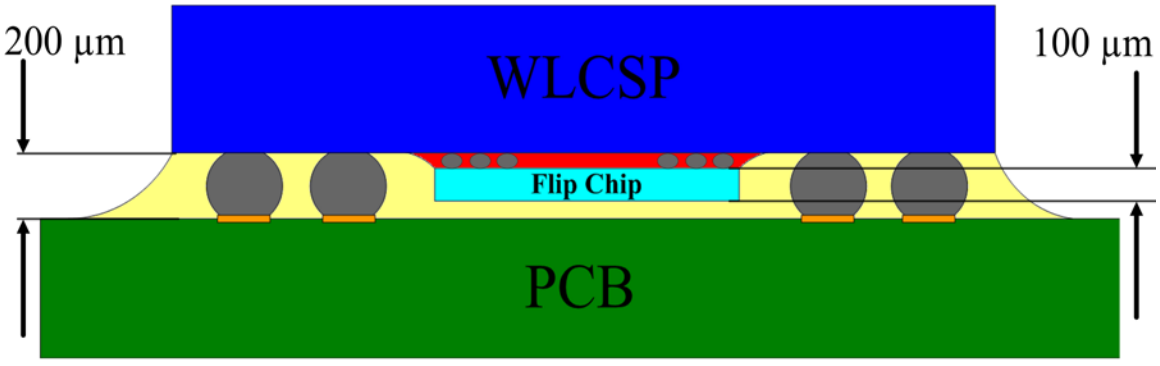


Figure 1-17 Schematic of a 3D-WLCSP Package Second Level Interconnection

Unlike the 3D structures discussed previously that stack dies or packages together with the wire-bonding method, the 3D-WLCSP technology utilizes the flip chip process to realize the thin profile, fine-pitch flip chip die to WLCSP integration on a wafer format. Also, in comparison with most of the 3D structures that stack the dies together

one on top of the other, the proposed 3D-WLCSP architecture realizes the 3D integration by face-to-face bonding of the active surface of the flip chip die to the active face of the WLCSP component. Due to the face-to-face bonding structure, both the signal transmission path and the electrical transmission path in the package can be greatly reduced so that the overall electrical performance of the package can be improved.

The face-to-face bonding structure of flip chip to WLCSP interconnection also serves as a good solution to reducing the package height. The overall package height is critical as the pursuit of smaller form factors continues in the consumer electronics market, especially in the field of handheld devices. In other words, the package height determines the usefulness and competitiveness of the packaged final product.

Furthermore, the utilization of flip chip process allows the manufacturing process of the 3D-WLCSP to be compatible with the conventional SMT process. The manufacturing cost per package is greatly reduced as the first-level assembly is accomplished at the wafer scale. Moreover, the test and burn-in of packages can also be performed at the wafer level so that “known-good-package” can be manufactured in such a fashion

CHAPTER 2 PROBLEM STATEMENT AND RESEARCH OBJECTIVES

2.1 Problem Statement

The 3D-WLCSP packaging structure leverages both low cost flip chip process and the existing infrastructures of high throughput wafer level packaging technology. However, there are several manufacturing challenges and processing difficulties toward the realization of this novel packaging technology.

The first challenge is the yield of both first level assembly and second level assembly. The first level assembly features a fine pitch Pb-free flip chip assembly process. And as the flip chip carries tiny solder bumps, it is highly challenging for the setup of the proper surface mount technology (SMT) process such as the wafer assembly placement program, reflow profile setting, choice of flux material, reflow ambient gas environment (level of oxygen inside the reflow oven), etc., in order to achieve a high assembly yield (>99%). The second level assembly utilizes the standard CSP assembly process to realize the chip to organic PCB interconnection. The choice of solder paste material as well as the setup of proper Pb-free reflow profile are critical to achieving high assembly yield and good solder joint quality.

The second challenge facing this novel 3D-WLCSP technology is the package long-term reliability performance, both at the first level and second level. A high reliability performance is always desired as it determines the endurance of the package in

the field application. Thus it is critical to find the proper process parameters and select the right process material, in order to achieve high package reliability performance.

The third challenge for the 3D-WLCSP is the manufacturing throughput. Factors such as the speed of pick-and-placement machine, placement program efficiency, placement machine vision recognition, underfill application and so on, will all affect the manufacturing throughput.

Therefore, it is vital to address these challenges discussed above and understand how the process parameters as well as material selection such as underfill and flux would affect the 3D-WLCSP process yield, reliability and manufacturing throughput so that a robust manufacturing process for the novel 3D-WLCSP can be achieved.

2.2 Research Objectives and Dissertation Structure

The main objective of this research is to characterize the manufacturing technology including process parameter settings, material selections as well as process qualifications, in order to achieve high yield, satisfactory reliability performance and high throughput on the 3D-WLCSP first-level and second-level packages. Five studies are designed to achieve the main research objective and are presented in the chapter 3 through chapter 7 as follows:

Chapter 3 explores the impact of reflow profile parameters for the 3D-WLCSP first-level assembly. This study focuses on the assembly yield, solder joint quality as well as the thermal reliability of the first-level packages assembled with fifteen reflow profiles. The solder joints formed by different reflow profiles are also examined by cross-sectional micrographs to understand the impact of the reflow profile parameters on

the intermetallic compound formation as well as the solder joint shape. Another purpose of this study is to understand the reflow process window, the reflow ambient gas environment as well as the flux amount applied, so that both high yield and satisfactory reliability performance can be achieved on the 3D die-to-wafer first-level assembly process.

Chapter 4 presents the flux and underfill compatibility study on the 3D-WLCSP first-level packages. The purpose of this study is to evaluate different flux/underfill candidate for the best performing material combination so that the package long-term reliability performance can be achieved.

Chapter 5 explores the no-flow underfill assembly process development for the 3D-WLCSP first-level packages. The purpose of this study is to explore an alternative method for the 3D die-to-wafer integration for a shorter processing time, lower cost and higher throughput.

Chapter 6 focuses on the first-level assembly scale up study. Effort is put on the creation and improvement of the flip chip wafer level assembly program for high processing yield and throughput.

Chapter 7 focuses on the second-level assembly process development including the solder paste selection as well as the reflow profile study. Thermal reliability testing is also performed to explore the reliability performance of the second level packages.

Chapter 8 summarizes the findings and conclusions of this research.

Chapter 9 describes the contributions of this research. Recommendations of future work are also presented in this chapter.

CHAPTER 3 PB-FREE REFLOW PROFILE STUDY FOR 3D-WLCSP FIRST-LEVEL ASSEMBLY

3.1 Introduction

The 3D-WLCSP first level assembly realizes the die-to-wafer integration by the flip chip process, as shown in Figure 3-1. First the die is picked by the die-attach machine and dipped into the flux tray. Then the machine will place the dies onto each individual WLCSP substrate fabricated on the same wafer. After the chip placement, the whole wafer is reflowed to form the first-level solder joint interconnects in each individual first-level packages on the wafer.

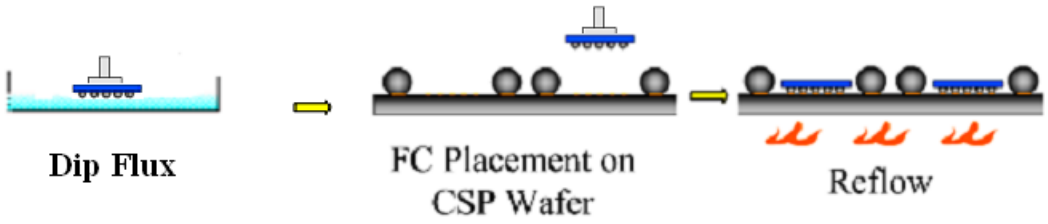


Figure 3-1 3D-WLCSP First Level Assembly Flip Chip Process

One advantage the flip chip process has over the traditional wire bonding technology is the compatibility with conventional surface mount technology (SMT) process, as the reflow is being utilized to attach the solder bumped flip chip to the substrate.

Reflow is a controlled heating process for the solder to melt and wet on the substrate pads so that the interconnections between die and substrate can be achieved. Figure 3-2 shows a typical reflow profile for Pb-free flip chip assembly.

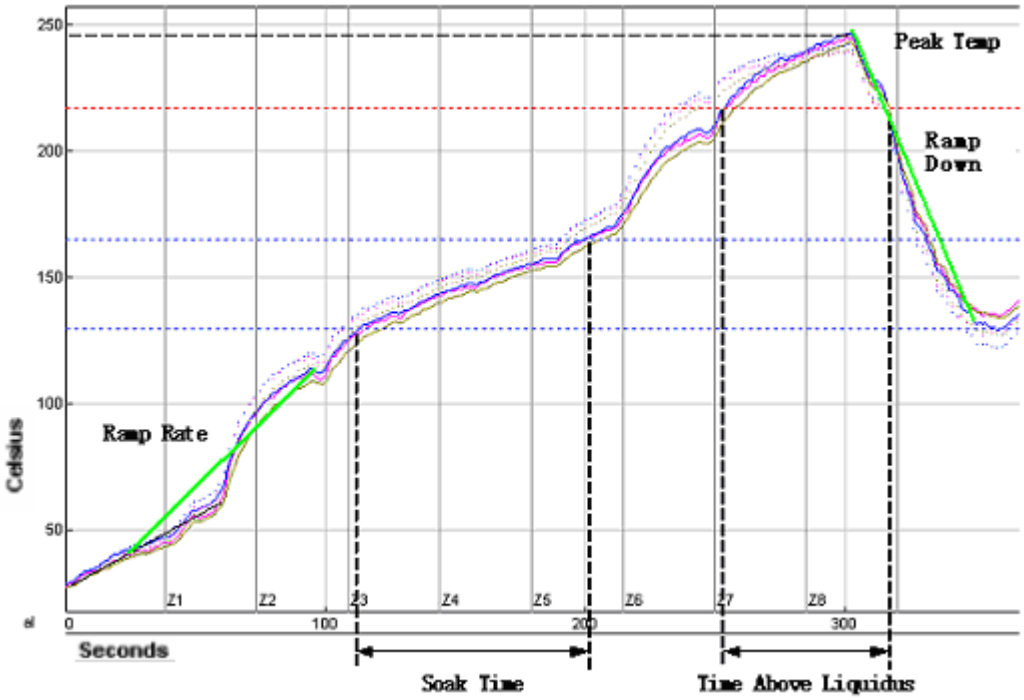


Figure 3-2 Example of Pb-free Reflow Profile

Several parameters in the reflow profile may affect the solder joint formation. These parameters include the ramp-up rate, soak time, peak temperature, time above liquidus, and ramp rate during cooling [35]. The ramp-up rate measures how fast the temperature increases on the reflowed sample at the heating stage. In comparison, the ramp-down rate depicts how quick the temperature drops on the reflowed sample at the cooling stage.

The purpose of the soak zone in the reflow profile is to activate the flux agent so that the oxidation on pads and solder bumps can be effectively removed. This activity

will prepare the surface for the solder to wet. The other purpose of soak zone is to make the temperature across the board equalized before the board enters the reflow zone. The soak time depicts the time that the reflowed component spends in the soak zone during reflow.

The length of time during which the solder is heated up to above the solder melting temperature (183°C for Eutectic Sn63/Pb37 solder, 217°C for SAC305 solder) is referred as “time above liquidus” (TAL). The maximum reflow temperature (peak temperature) is achieved during the TAL.

Harrison [36] discussed the Pb-free reflow soldering of SnAg3.8Cu0.7 alloy for QFP packages. The author did a reflow process windows study which shows that a good Pb-free SnAg3.8Cu0.7 solder joint can be formed with a peak temperature as low as 225°C.

Salam [37] studied the Pb-free reflow profile for the solder bumps on FR4 substrate. It was reported that the most significant factor in achieving a joint with a thin intermetallic compound (IMC) layer and fine microstructure is the peak temperature.

Pan [38] studied the effect of the reflow peak temperature and time above liquidus (TAL) on both SnPb and SnAgCu Pb-free solder joint shear strength with 4 different chip resistors on FR4 substrate. The author conducted a full-factorial experiment on three levels of peak temperature and three levels of TAL. It was confirmed by the author that both reflow peak temperature and TAL of Pb-free reflow profiles are critical factors in terms of the package shear strength of the SnAgCu solder joints.

Ladani [39] presents the approach to study the effect of several manufacturing variables on the product defects as well as on the durability of BGA Pb-free solder joints.

The paper found that inadequate reflow profile produced insufficient wetting and insufficient intermetallic formation. The author conducted experiment and statistical analysis to find out that heating ramp and peak temperature have nonlinear effects on the thermal cycling durability.

P. L. Tu and Y. C. Chan [40] discussed the vibration fatigue failure of μBGA solder joints formed with different reflow profiles. The author found that the initial formation of the IMC layer during reflow ensures a good metallurgical bond but that a thicker IMC layer results in a shorter vibration fatigue lifetime.

Other publications also discussed the solder joint failures during the thermal cycling reliability test, where the solder joints are exposed to high thermal stress. The thermal stress caused fatigue on solder interconnection, a commonly observed failure mode [41]. The failure mechanisms were found to be inadequate wetting or Cu–Sn intermetallic compound (IMC) layer fatigue between the solder and the Cu-pad [41,42]. The IMC layer has drawn a lot of attention in the electronic packaging industry with Pb-free application [40, 42]. In general, an increase of IMC layer thickness decreases the lifetime of solder joints, as the intermetallic is the most fragile part in the solder joint. Typically the IMC layer thickness strongly depends on reflow process conditions such as time above liquidus and peak temperature [41-44]. The correlation between reflow process and the IMC layer was investigated to find an ideal reflow process condition, subsequently achieving reliable solder joints in thermal cycling test using statistical analysis [45].

Although there are many previous publications on the effect of reflow profile toward Pb-free solder joint formation and reliability, none of them conducted study based

on the fine-pitch flip chip on silicon substrate assembly. The 3D-WLCSP at first-level is a face-to-face, silicon on silicon bonding architecture. The CTE of the silicon device is ~3 ppm/°C while the FR4 organic substrate has a CTE of ~18 ppm/°C. Compared with the silicon to organic substrate assembly, the flip chip to WLCSP assembly has no CTE mismatch between the die and substrate as they are both made of silicon. Thus the thermal reliability performance of the silicon-on-silicon packages can be quite different from that of the flip chip on PCB packages.

In the meanwhile, the silicon-on-silicon packaging structure is getting more and more industry attention as it is seen often in the booming 3D packaging technologies. However, compared with the traditional flip chip on organic PCB assembly, very little research has been done in this packaging structure, especially in the silicon-on-silicon assembly with flip chip process. As the reflow profile is the key process for the solder joint formation of the silicon-on-silicon flip chip assembly, it is important to understand how the reflow process parameters will impact the package yield and solder joint formation as well as package thermal reliability performance based on the silicon-on-silicon flip chip assembly.

In this chapter, a comprehensive study on the impact of reflow profile parameters toward the 3D-WLCSP first-level silicon-on-silicon assembly is carried out. The focus is put on three reflow profile parameters - soak time, peak temperature and time above liquidus. Assembly yield was calculated based on each reflow process condition with different reflow parameter settings to find out how the process parameters would affect the assembly yield. Both the nitrogen reflow environment and the air reflow environment were investigated to learn the impact of the reflow ambient gas on the assembly yield as

well as the solder joint quality. Cross section was performed to reveal the intermetallic compound formation under each process condition. Shear test was also performed to understand how the silicon-on-silicon package solder joint strength is affected by the reflow profile conditions. A liquid-to-liquid thermal shock (LLTS) reliability test was performed to evaluate the silicon-on-silicon packages assembled with different reflow profiles.

3.2 Experimental Approach

3.2.1 Test Vehicle Design and Fabrication

Four test vehicle designs fabricated by FlipChip International (FCI) were utilized in this work. The four test vehicles were simultaneously manufactured on the same wafer by dividing the WLCSP carrier substrate and complimentary flip chip test wafer into quadrants (listed as Quadrant 1 to Quadrant 4 in Figures 3-3 and 3-4). The test vehicle wafer containing four quadrants of WLCSP substrates used in this research is shown in Figure 3-5. The design of the WLCSP and the flip chip test vehicles must support the development requirements of the 3D-WLCSP manufacturing technology. As part of this effort, daisy chain test vehicles were produced so that the electrical continuity can be monitored across the daisy chain on the 3D packages to understand the assembly yield as well as the package long-term reliability performance during environmental testing. The four test vehicles have different pitches, bump sizes and layouts, which are designed to allow researchers to compare and study the effects of these design features on different field applications.

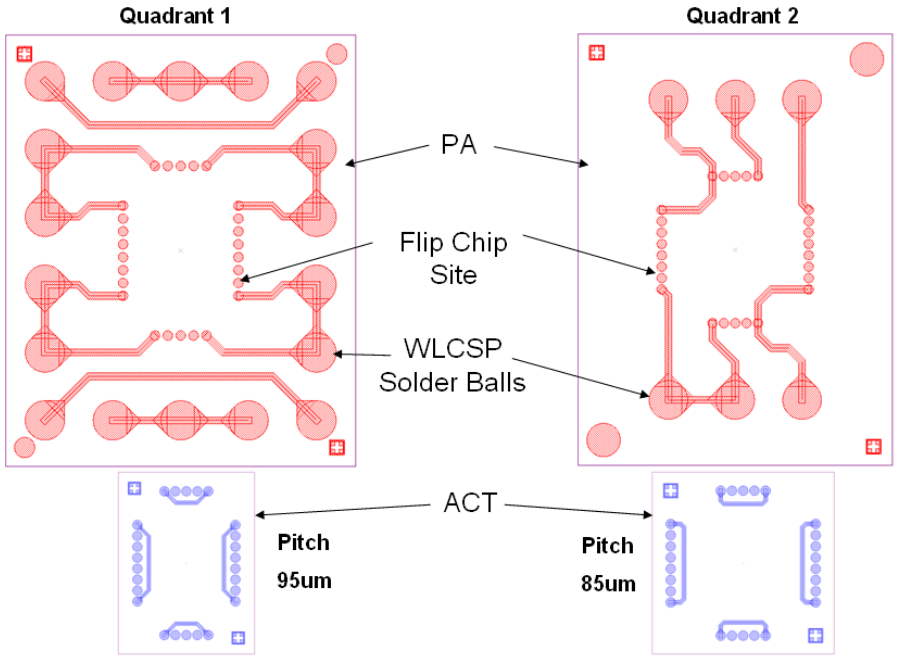


Figure 3-3 Quadrant 1 and Quadrant 2 Test Vehicle

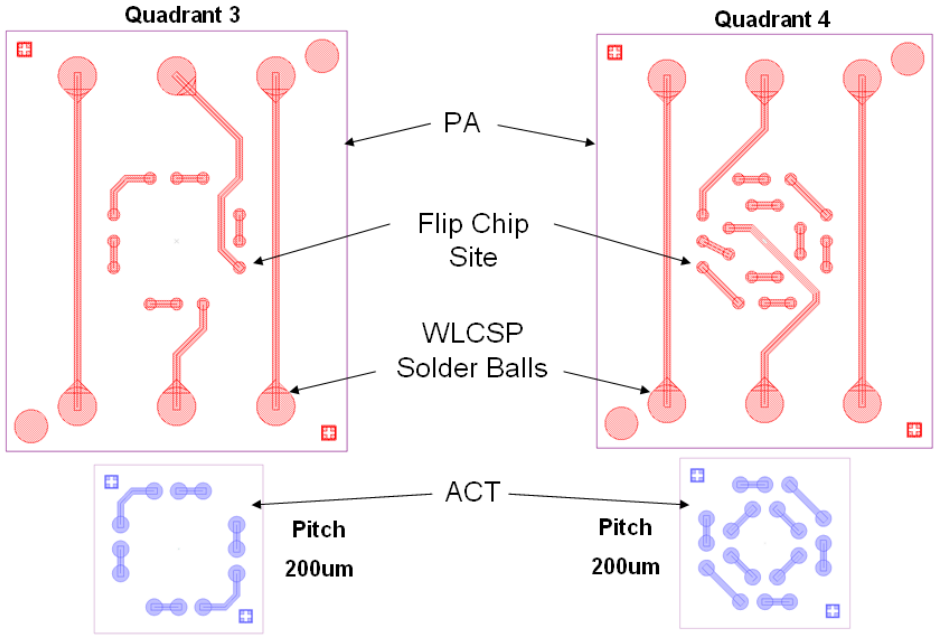


Figure 3-4 Quadrant 3 and Quadrant 4 Test Vehicle

The Passive Component (PA) device forms the base substrate for the 3D first-level package and is a solder balled wafer level CSP in wafer form, as shown in Figure 3-5. The active component (ACT) dies are face-to-face mounted to the PA wafer by the flip chip process. All of the devices used for this work contain Pb-free SAC 305 (96.5% tin, 3% silver, 0.5% copper) solder bumps.

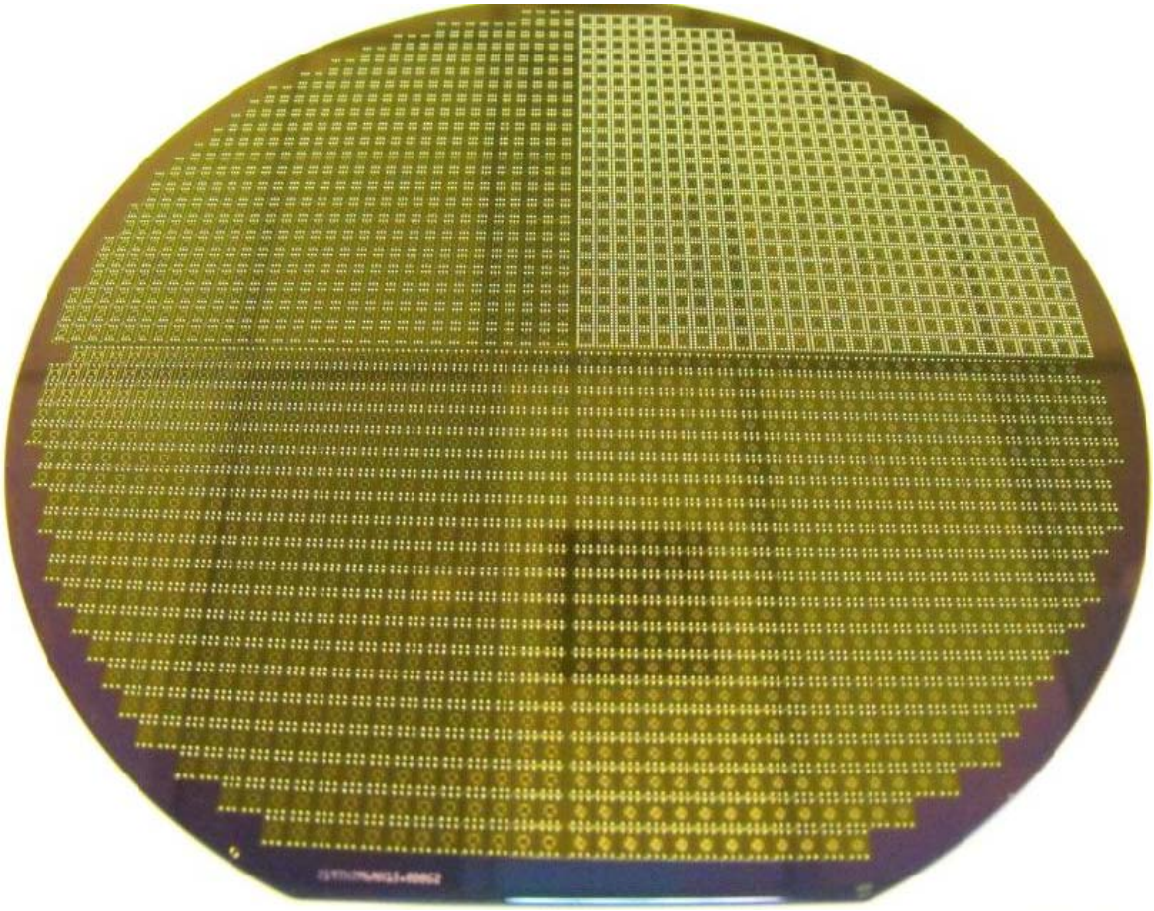


Figure 3-5 3D WLCSP PA Test Vehicle Wafer

The ACT die from Quadrant 1 has a 95-micron pitch with 26 perimeter bumps per component. The dimension of Quadrant 1 flip chip die outlines 1.1mm x 1.5mm. Figure

3-6 shows the image of a Quadrant 1 ACT die taken by a scanning electron microscope (SEM). Figure 3-7 shows a close-up view of two adjacent Quadrant 1 bumps featuring 95-micron pitch and bump height of only 46 microns. The base PA device for Quadrant 1 is 2.5mm by 3.1mm with 18 perimeter bumps with a 0.5 mm pitch and 0.27 mm metallization as shown in Figure 3-8.

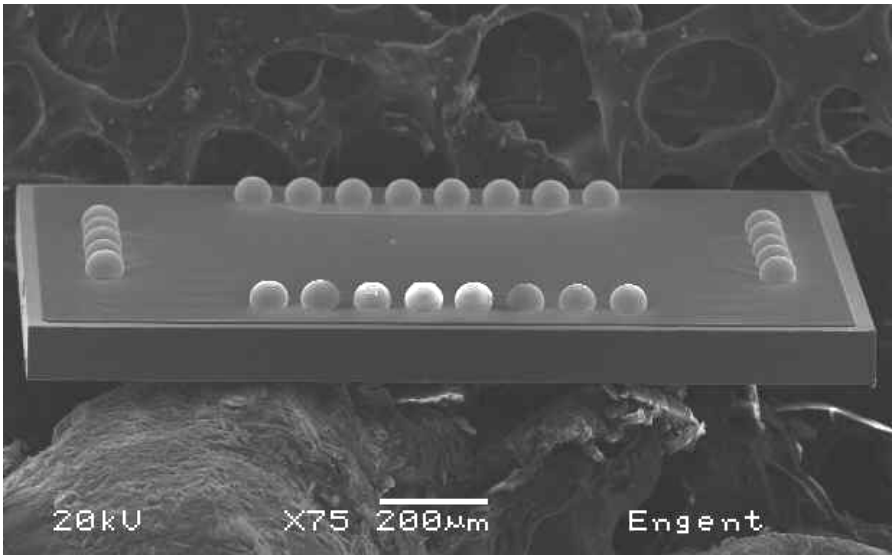


Figure 3-6 Quadrant 1 Test Vehicle ACT SEM Image

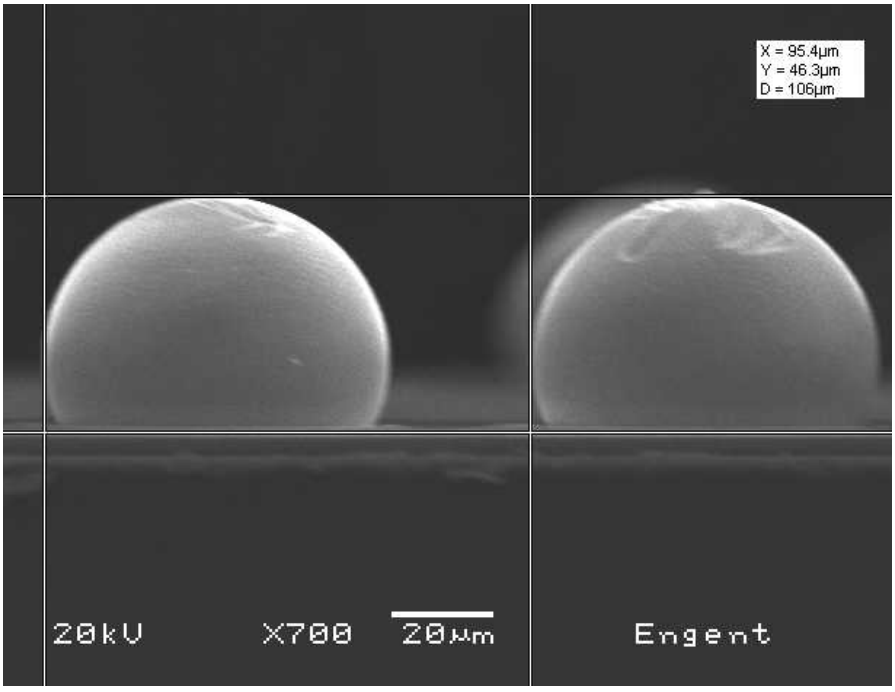


Figure 3-7 Quadrant 1 Test Vehicle ACT Adjacent Solder Bumps

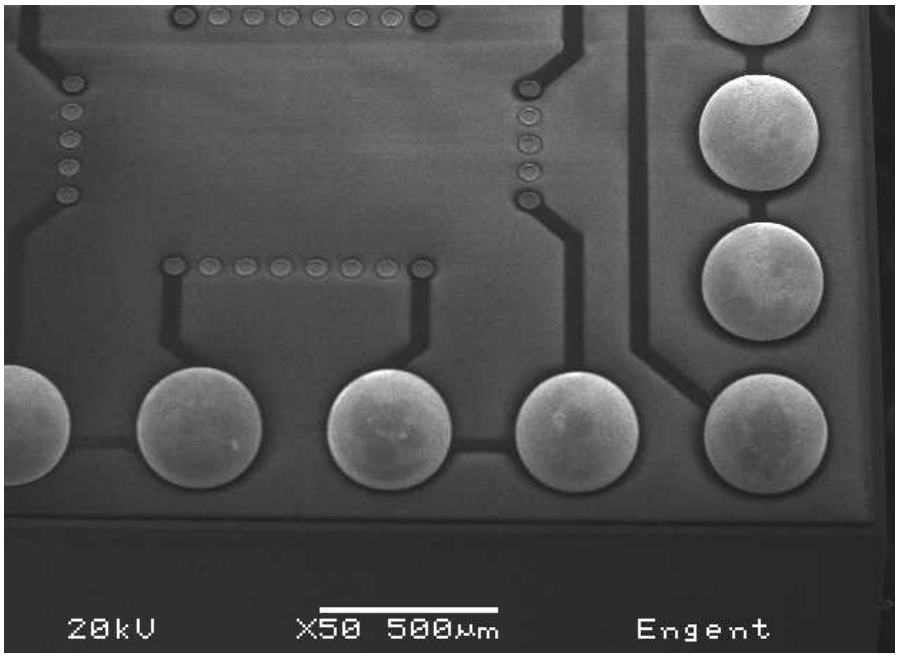


Figure 3-8 Quadrant 1 PA SEM Image

The die from Quadrant 2 has an 85-micron pitch with 26 perimeter bumps per component. The base PA substrate device for Quadrant 2 has 6 bumps (3 per short side) at a 0.4 mm pitch with 0.27 mm metallization as shown in Figure 3-3. The die from Quadrant 3 has a 200-micron pitch with 12 perimeter bumps. The die from Quadrant 4 has a 200-micron staggered full array pitch with 20 bumps, as shown in Figure 3-9 and Figure 3-10. The base PA devices for Quadrant 3 and Quadrant 4 have 6 bumps (3 per short side) at a 0.75 mm pitch with 0.29 mm metallization for the pads. The details of the ACT and PA test vehicles are outlined in Table 3-1 and Table 3-2.

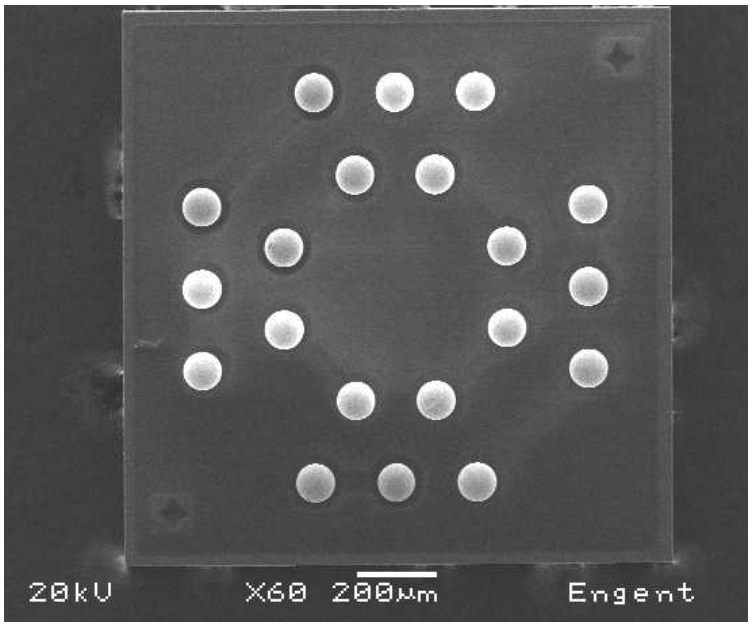


Figure 3-9 Quadrant 4 ACT SEM Image

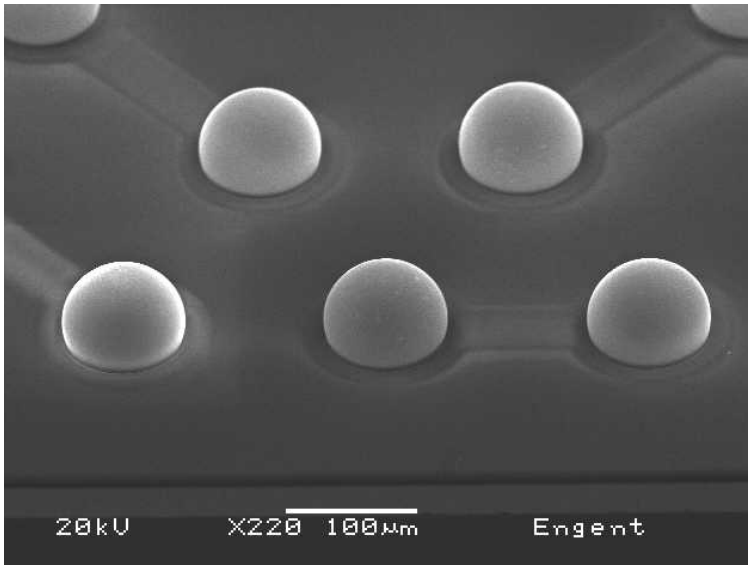


Figure 3-10 Quadrant 4 ACT solder bump SEM image

Table 3-1 Flip chip ACT Test Vehicle Information

Quadrant	Pitch/Bump Diameter (um)	Flip Chip I/O count	Flip Chip I/O array	UBM Diameter (um)	Die Size (um)
1	95/70	26	Peripheral	55	1100 x 1500
2	85/63	26	Peripheral	53	1300 x 1300
3	200/105	12	Peripheral	100	1300 x 1300
4	200/95	20	Full	90	1300 x 1300

Table 3-2 CSP PA Test Vehicle Information

Quad	CSP Pitch (um)	I/O Count	CSP I/O Array	UBM Diameter (um)	Die Size (um)
1	500	18	Peripheral	300	2500 x 3100
2	400	6	Peripheral	300	2270 x 3150
3	750	6	Peripheral	300	2500 x 3100
4	750	6	Peripheral	300	2500 x 3100

The fabrication of the test vehicles in this study leveraged FCI's SpheronTM RDL WLCSP technology [4] and Standard Flip ChipTM (SFCTM) with repassivation flip chip bumping technology [5]. The manufacturing process flows for each package element is

shown schematically in Figure 3-11. The primary design considerations for this type of package are overall package dimensions including overall package height, WLCSP stand off and flip chip bumping pitches. These considerations drive lower level design and manufacturing considerations such as individual package height, flip chip die thickness, required redistribution line/space width, and choice of assembly approaches including underfill applications.

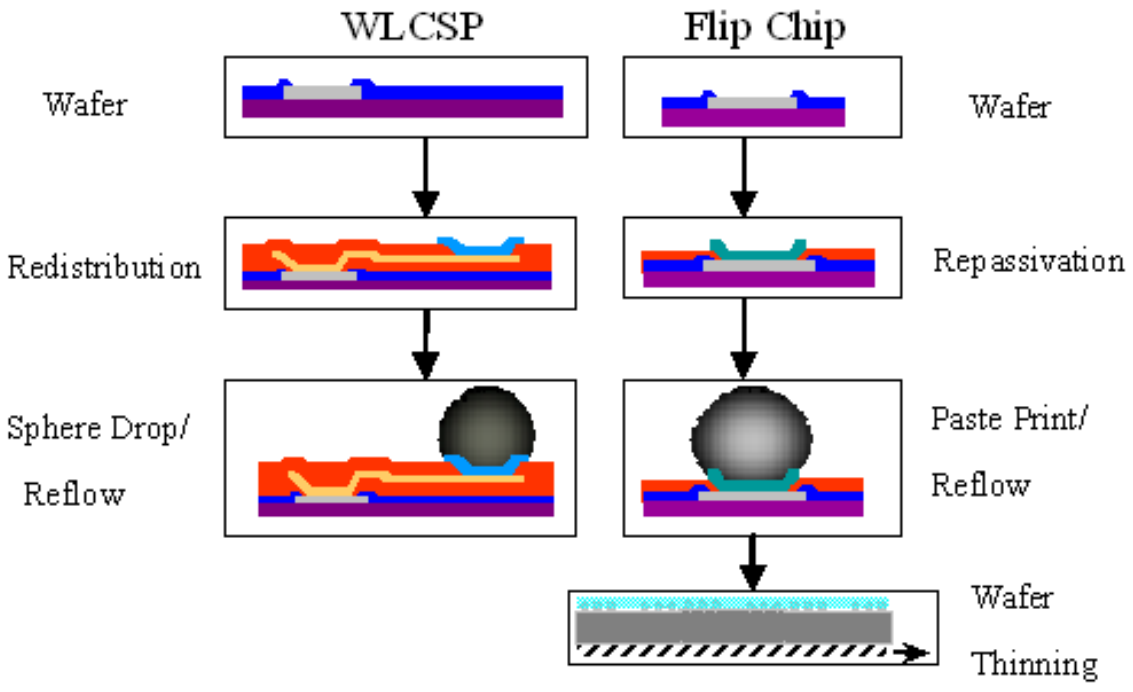


Figure 3-11 Fabrication Process of WLCSP and Flip Chip Test Vehicles

The peripheral I/O layout for three of four ACT test samples was chosen to represent a typical ASIC driver I/O layout initially intended for wire-bonding interconnection. To obtain the lowest cost bumping structure for this type of I/O layout, a repassivated flip chip structure was maintained. This bump on the I/O structure was designed to provide a low bump height (<50 microns) and minimal post assembly

stand-off. The resultant low profile standoff of this chip-to-chip interconnection supports the ultimate stand-off requirements of the finished 3D-WLCSP first-level package. Flip chip technologies based on advanced paste printing approaches like FCI's proprietary SFC^{TM} technology can readily support sub-100 micron pitch flip chip bumping applications [32-34] and were successfully used in the fabrication of the ACT test vehicles for pitches ranging between 85 and 200 microns.

For the test vehicles in this study, the size of solder balls on PA WLCSP is 300 microns which creates an approximate 200 microns stand off from a printed circuit board surface, depending on the PCB pad size. Package height dictates the usefulness of the end product in its intended application. The die-to-wafer integration structure by the 3D-WLCSP technology serves the need to meet this height limit and reduce required printed circuit board space by face-to-face bonding the active flip chip component within the footprint of a larger WLCSP component. The face-to-face bonding structure requires that the flip chip have a thin profile of about 100 microns so that the backside of the flip chip does not contact the PCB surface in the second-level assembly. This indicates both a low stand off in the flip chip bumping structure and a thin profile of the flip chip die at the height of about 100 microns, which can be achieved by thinning the flip chip in the wafer format. The wafer thinning process also requires stress relief after grinding by either mechanical polishing, chemical mechanical polishing (CMP), or wet etch to maintain wafer strength, chip integrity and best possible die surface flatness for later manufacturing processes.

3.2.2 DOE Layout for Reflow Profile Study

A design of experiment (DOE) utilizing response surface methodology (RSM) is conducted in this study. The response surface methodology is a statistical approach for the analysis of problems in which a response of interests is impacted by several variables [46]. The RSM can explore the relationship between the response and variables, as well as the quadratic effects of each variable.

Another advantage that a RSM design has over the full-factorial design is the reduction on the number of experiments. In this study, the factors investigated include the reflow profile soak time, peak temperature and time above liquidus. With a full-factorial design matrix in which three levels of each factor is studied, a total of 27 (3x3x3) experiments should be conducted. In comparison, an RSM design with two center points takes only 16 runs.

In this study, the three levels of soak time investigated are at 40 seconds, 80 seconds, and 120 seconds, respectively. The three levels for the reflow peak temperature were defined as 225 °C, 245 °C, and 265 °C. The melting temperature of SAC305 solder is 217 °C. For this reason, the 225 °C was picked as the low-level of the factor of peak temperature to ensure the solder can reach its melting temperature during reflow. The three levels of time above liquidus are 20 seconds, 80 seconds, and 140 seconds. A matrix of RSM central composite design is shown in Table 3-3. In the design pattern, the first number stands for the level of soak time, the second number represents the level of peak temperature, and the third number indicates the level of the time above liquidus (TAL). For example, the experimental pattern of (+1 –1 –1) means a reflow profile with a soak time of 120 seconds (high level), peak temperature of 225 °C (low level) and a TAL

of 20 seconds (low level). The response of the design matrix is the experiment results based on different reflow profiles. In this study, package shear test strength, intermetallic compound characteristics, package yield and reliability performance are the responses that are to be investigated. Two center points (0,0,0) were put in the design matrix. The one added center point was intended to provide an unbiased estimate of the variance.

Table 3-3 Response Surface Methodology DOE Layout

Pattern	Soak time (s)	Peak Temp (°C)	TAL (s)	Response
+1-1-1	120	225	20	
+1+1+1	120	265	140	
+1 0 0	120	245	80	
0 0 0	80	245	80	
-1 0 0	40	245	80	
0 0-1	80	245	20	
0 0+1	80	245	140	
+1-1+1	120	225	140	
-1+1-1	40	265	20	
-1-1-1	40	225	20	
0+1 0	80	265	80	
+1+1-1	120	265	20	
-1+1+1	40	265	140	
0 0 0	80	245	80	
-1-1+1	40	225	140	
0 -1 0	80	225	80	

3.2.3 Reflow Profiles Setup

Based on Table 3-3, fifteen reflow profiles are configured with a KIC profiler and the Rehm RN38 convection reflow oven. Two thermal couples were attached to the test vehicle profiled, in order to make sure the readings are identical and correct. Figure 3-12 shows the condition of (-1 -1 -1) in Table 3-3, which represents a soak time of 40

seconds, reflow peak temperature of 225 °C and the TAL of 20 seconds. Figure 3-13 shows the configuration of (+1 +1 +1), which has the soak time of 120 seconds, peak temperature of 265 °C and a TAL of 140 seconds. Each reflow profile was tested twice by the KIC profiler, in order to make sure the reflow profile is accurate and repeatable.

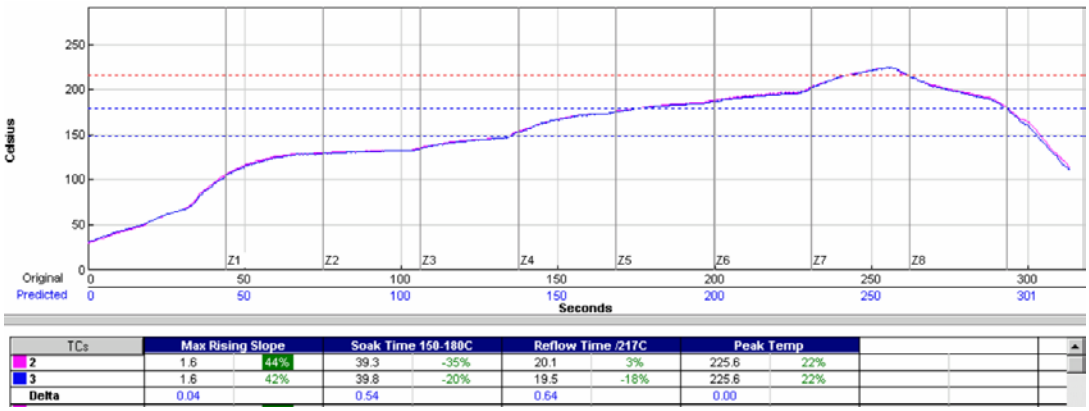


Figure 3-12 Reflow Profile Condition (-1 -1 -1)

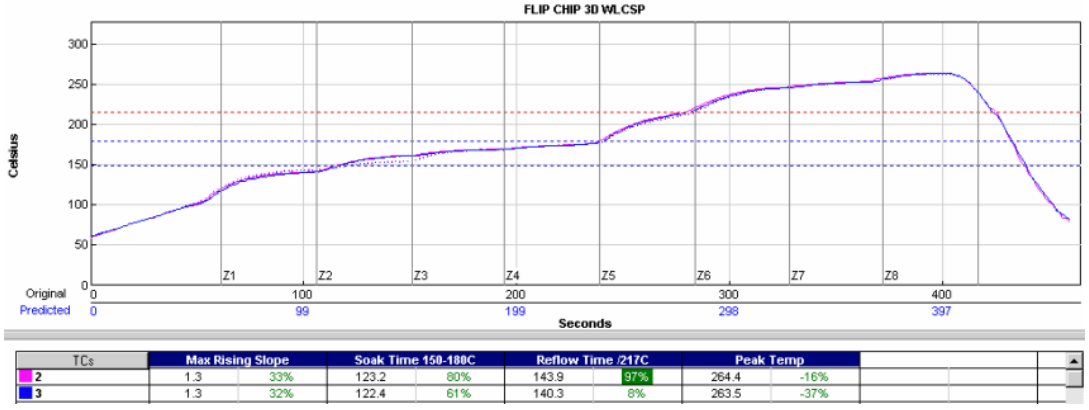


Figure 3-13 Reflow Profile Condition (+1 +1 +1)

3.3 Experimental Results and Discussions

3.3.1 Impact of Reflow Profiles on Package Yield with Nitrogen Environment

Both Quadrant 1 and Quadrant 3 test vehicles were used in this study. The 2x2 and 3x3 WLCSP tiles were used for the development work. The tiles were diced off from the WLCSP test vehicle wafer shown in Figure 3-5. The purpose of using the tiles instead of the whole wafer is to save the test vehicles. One Quadrant 1 2x2 WLCSP tile is shown in Figure 3-14.

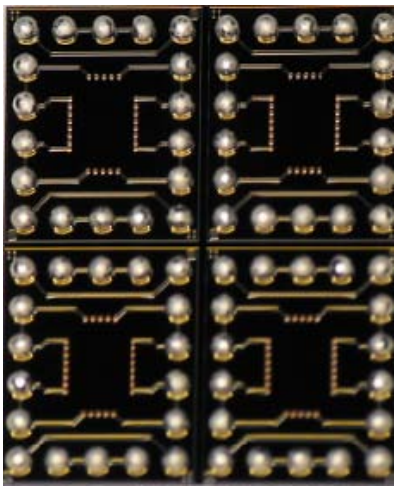


Figure 3-14 Quadrant 1 2x2 Test Vehicle Tile

A total of 480 samples of Quadrant 1 and 480 samples of Quadrant 3 first-level packages were assembled with 30 samples assembled per group according to Table 3-3. The assembly follows a standard flip chip assembly process, in which the flip chip is picked, dipped in a flux reservoir, vision aligned, and placed onto the substrate, as shown in Figure 3-1. The bump height of the Quadrant 1 flip chip die is 45 microns and the bump height of Quadrant 3 flip chip die is 70 microns. The flux dip height was 30

microns for Quadrant 1 assembly and 45 microns for Quadrant 3 assembly, as the Quadrant 3 flip chip test vehicles have larger solder bumps than those of Quadrant 1 flip chip test vehicles. After the chip placement, the samples were sent to the convection reflow oven with nitrogen environment following respective reflow conditions shown in Table 3-3. The yield was calculated based on the measurement of resistance across the daisy chain on each assembled package.

No failure was observed by the daisy chain probing measurement on any of the 900 assemblies. The result as listed on Table 3-4 shows that 100% yield was achieved on samples built under each of the fifteen reflow profile conditions, for both Q1 and Q3 test vehicles. This also indicates a wide reflow process window for the studied 3D-WLCSP first level silicon-on-silicon packages. Figure 3-15 shows the cross section images of the solder joint reflowed under different conditions. The left image of Figure 3-15 is the solder joint reflowed at (+1 +1 +1) reflow condition, which represents 120 seconds soak time, 265 °C peak temperature and 140 seconds reflow time (time above liquidus). The right image of Figure 3-15 shows the solder joint formed under the profile (-1 +1 -1), which features a soak time of 40 seconds, peak temperature of 265 °C and the reflow time of 20 seconds. The solder joint cross sections shown in Figure 3-15 as well as cross sections performed on samples built under the other thirteen reflow profiles, all show good solder joint quality. These cross sections confirm the 100% yield on the samples assembled under each reflow condition.

Based on the results shown in Table 3-4, it can be seen that the assembly results are insensitive to the reflow profile parameters with nitrogen reflow environment, which indicates a wide range of assembly window for the fine-pitch flip chip to silicon

assembly. This result also indicates that as long as the reflow profile parameters are within the process window of design matrix shown in Table 1, it can be estimated that the assembly will achieve 100% yield.

Table 3-4 3D-WLCSP First-level Assembly Yield Vs. Reflow Profile Condition

Pattern	Soak time (s)	Peak Temp (°C)	TAL (s)	Q1 Yield	Q3 Yield
+1-1-1	120	225	20	30/30	30/30
+1+1+1	120	265	140	30/30	30/30
+1 0 0	120	245	80	30/30	30/30
0 0 0	80	245	80	30/30	30/30
-1 0 0	40	245	80	30/30	30/30
0 0-1	80	245	20	30/30	30/30
0 0+1	80	245	140	30/30	30/30
+1-1+1	120	225	140	30/30	30/30
-1+1-1	40	265	20	30/30	30/30
-1-1-1	40	225	20	30/30	30/30
0+1 0	80	265	80	30/30	30/30
+1+1-1	120	265	20	30/30	30/30
-1+1+1	40	265	140	30/30	30/30
0 0 0	80	245	80	30/30	30/30
-1-1+1	40	225	140	30/30	30/30
0 -1 0	80	225	80	30/30	30/30

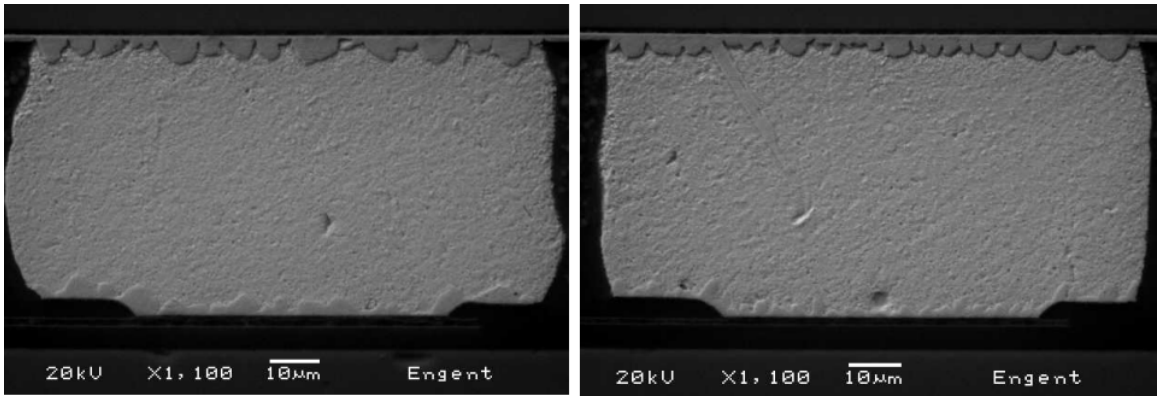


Figure 3-15 Cross Section of Quadrant 3 Solder Joint Reflowed at Condition (+1 +1 +1) (left) and (-1 +1 -1) (right)

3.3.2 Impact of Reflow Profiles on Intermetallic Compound Thickness

From Figure 3-15, it can be seen that the Quadrant 3 solder joints formed with reflow condition (+1 +1 +1) and (-1 +1 -1) show different intermetallic compound (IMC) thickness.

Figure 3-16 shows the close up view at the IMC layer of the solder joint formed with the reflow profile (+1 +1 +1). The highest peak of the IMC layer is measured as 4.6µm in the cross section.

Figure 3-17 shows the close up view at the IMC of the solder joint formed with the reflow profile (-1 +1 -1) with a highest peak of 3.73µm.

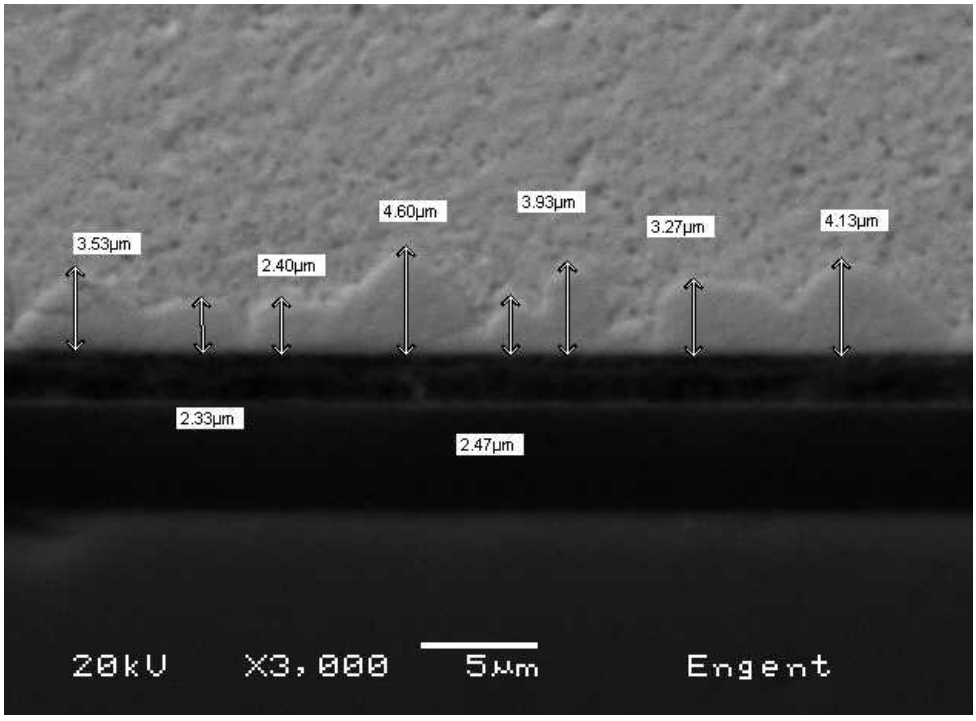


Figure 3-16 Close up View at the IMC of Reflow Profile Condition (+1 +1 +1)

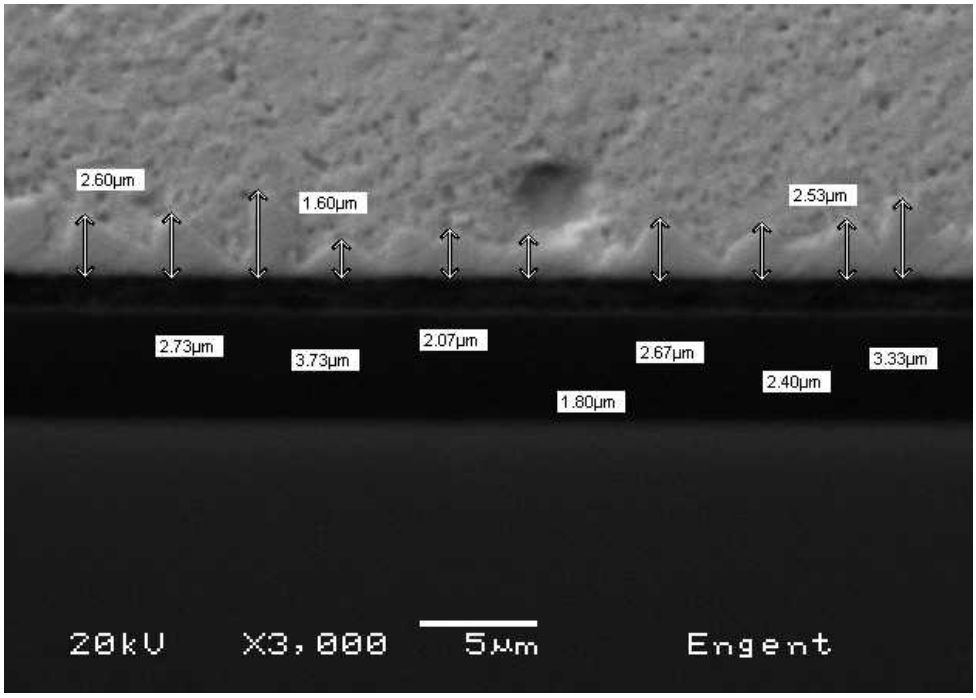


Figure 3-17 Close up View at the IMC of Reflow Profile Condition (-1 +1 -1)

In order to quantify the solder joint IMC layer thickness on packages assembled under each reflow profile condition, a mathematical method is carried out with the Simpson's rule, as shown in Equation 3-1. The total IMC area, as shown in Figure 3-16, is composed of several peak areas. Simpson's rule is a method of integration approximation and is used to numerically calculate the area of each individual peak area. Figure 3-18 shows the example of one calculation of the IMC with Simpson's rule. For each peak area, the heights of start, middle and end locations were measured, with the help of SEM and the image processing method.

$$\int_a^b f(x)dx \approx \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

Equation 3-1 Simpson's Rule on Area Calculation

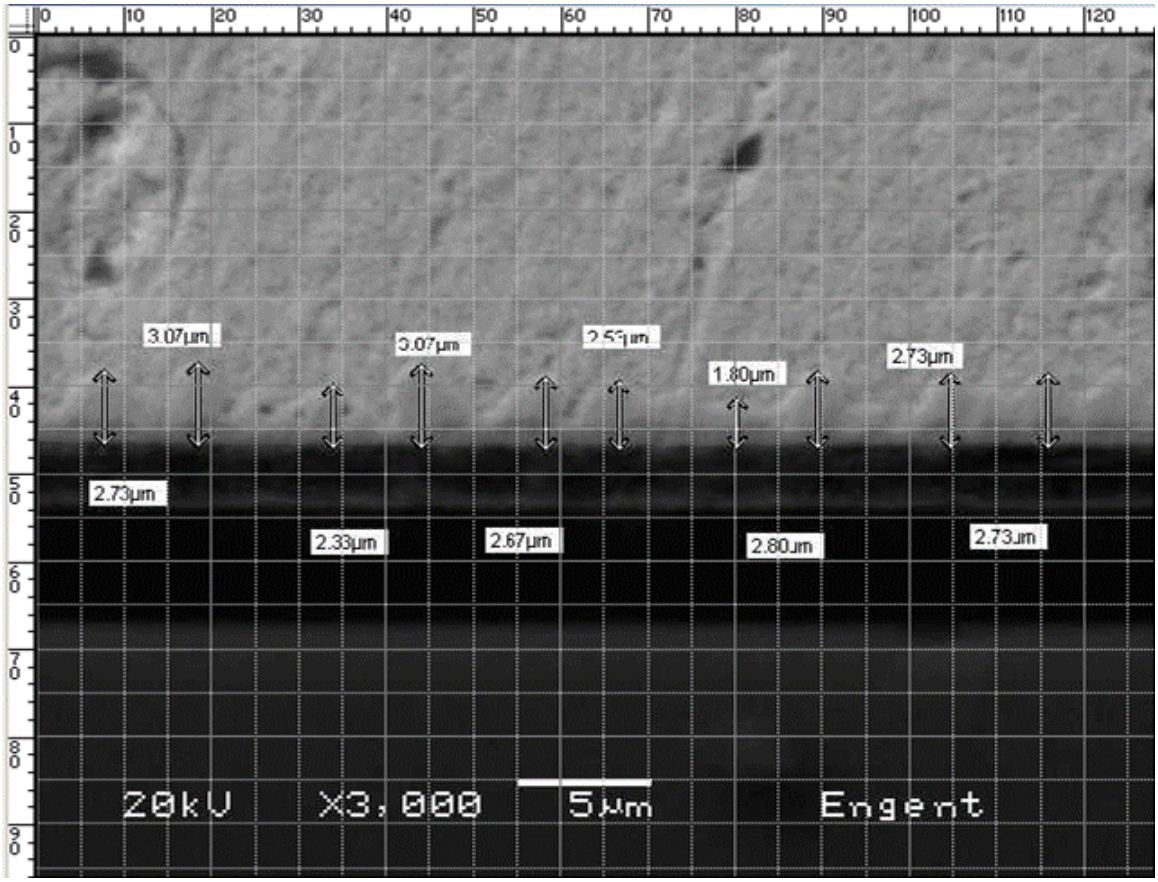


Figure 3-18 Calculation of the IMC by Simpson's Rule

The total IMC area for each cross section was obtained by adding the IMC area of each individual peak area calculated by the Simpson's rule. The average IMC thickness is calculated by dividing the total IMC area by the length of IMC studied, which was measured by the SEM. Three cross section samples from each reflow profile condition are used to calculate the IMC area. Then the average IMC thickness of each reflow condition is obtained by averaging the three studied samples. By this means, the average solder joint IMC thickness of samples reflowed under each reflow condition is calculated and shown in Table 3-5.

Table 3-5 IMC Thickness Vs. Reflow Profile Condition

Pattern	Soak time (s)	Peak Temp (°C)	TAL (s)	Average IMC Thickness (μm)
+1-1-1	120	225	20	1.46
+1+1+1	120	265	140	2.49
+1 0 0	120	245	80	2.23
0 0 0	80	245	80	2.18
-1 0 0	40	245	80	2.35
0 0-1	80	245	20	1.19
0 0+1	80	245	140	2.70
+1-1+1	120	225	140	2.45
-1+1-1	40	265	20	1.82
-1-1-1	40	225	20	1.43
0+1 0	80	265	80	2.22
+1+1-1	120	265	20	1.83
-1+1+1	40	265	140	2.42
0 0 0	80	245	80	2.32
-1-1+1	40	225	140	2.43
0 -1 0	80	225	80	2.26

The average IMC thickness data was input to JMP software in the response surface design matrix. The data was then analyzed by JMP with results shown in Figure 3-19. Based on the statistical analysis, it can be easily seen that the time above liquidus is the most significant factor for the IMC thickness, with a confidence level of more than 99.9%. The other two reflow profile parameters, soak time and peak temperature, were not significant for the intermetallic compound thickness. None of the interactions among factors and quadratic effects were found to be significant at a 95% level. Based on the results shown on Table 3-5, it can be found that higher time above liquidus leads to thicker IMC thickness. This result indicates that for the 3D-WLCSP first level silicon-on-silicon assembly, in order to control the IMC thickness of the solder joint, the only factor needs to be controlled is the time above liquidus during the reflow process.

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Soak Time(40,120)&RS	1	1	0.0000100	0.0003	0.9878
Peak Temp(225,265)&RS	1	1	0.0562500	1.4175	0.2788
Reflow Time(20,140)&RS	1	1	2.2657600	57.0982	0.0003
Soak Time(40,120)*Peak Temp(225,265)	1	1	0.0001125	0.0028	0.9593
Soak Time(40,120)*Reflow Time(20,140)	1	1	0.0003125	0.0079	0.9322
Peak Temp(225,265)*Reflow Time(20,140)	1	1	0.0666125	1.6787	0.2427
Soak Time(40,120)*Soak Time(40,120)	1	1	0.0119813	0.3019	0.6025
Peak Temp(225,265)*Peak Temp(225,265)	1	1	0.0007995	0.0201	0.8918
Reflow Time(20,140)*Reflow Time(20,140)	1	1	0.2031426	5.1193	0.0643

Figure 3-19 Analysis of Reflow Parameters on IMC Thickness

3.3.3 Impact of Reflow Profiles on Package Shear Strength

Shear test was performed on 18 Quadrant 4 assemblies under each reflow profile condition shown in Table 3-3. The shear results for each reflow condition were then averaged and filled in Table 3-6.

The data was analyzed by JMP software with the results shown in Figure 3-20. Based on the result, none of the three primary factors is significant in terms of the package shear strength. However, it was found that statistically the interaction between peak temperature and reflow time, and the quadratic effect of the reflow time are significant at the 95% level.

Table 3-6 Quadrant 4 Shear Strength Vs. Reflow Profile Condition

Layout	Soak Time (s)	Peak Temp(°C)	TAL (s)	Shear (kg)
-1-1-1	40	225	20	0.827
-1+1+1	40	265	140	0.906
-1-1+1	40	225	140	0.907
000	80	245	80	0.883
+100	120	245	80	0.845
00+1	80	245	140	0.926
+1+1+1	120	265	140	0.889
-100	40	245	80	0.856
+1-1-1	120	225	20	0.881
+1-1+1	120	225	140	0.924
0+10	80	265	80	0.826
-1+1-1	40	265	20	0.964
+1+1-1	120	265	20	0.973
00-1	80	245	20	0.979
0-10	80	225	80	0.842
000	80	245	80	0.892

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
X1(40,120)&RS	1	1	0.00027040	0.4595	0.5231
X2(225,265)&RS	1	1	0.00313290	5.3244	0.0605
X3(20,140)&RS	1	1	0.00051840	0.8810	0.3842
X1(40,120)*X2(225,265)	1	1	0.00078013	1.3258	0.2934
X1(40,120)*X3(20,140)	1	1	0.00049613	0.8432	0.3939
X2(225,265)*X3(20,140)	1	1	0.00877813	14.9185	0.0083
X1(40,120)*X1(40,120)	1	1	0.00081696	1.3884	0.2833
X2(225,265)*X2(225,265)	1	1	0.00306621	5.2110	0.0626
X3(20,140)*X3(20,140)	1	1	0.01877823	31.9137	0.0013

Figure 3-20 Analysis of Reflow Parameters on Package Shear Strength

In order to understand if there is any other reflow process parameter affecting the package shear strength, the cooling down rate of each reflow profile was measured by the

KIC profiler software. The reflow profile condition, the shear strength and the cooling down rate are shown in Table 3-7.

Table 3-7 Quadrant 4 Shear Strength Vs. Reflow Profile Cool Down Rate

Layout	Soak Time (s)	Peak Temp(°C)	TAL (s)	Shear (kg)	Cool Down Rate
-1-1-1	40	225	20	0.827	3
-1+1+1	40	265	140	0.906	4.35
-1-1+1	40	225	140	0.907	3.5
000	80	245	80	0.883	3.4
+100	120	245	80	0.845	3.35
00+1	80	245	140	0.926	3.65
+1+1+1	120	265	140	0.889	4.15
-100	40	245	80	0.856	3.65
+1-1-1	120	225	20	0.881	3.3
+1-1+1	120	225	140	0.924	3.45
0+10	80	265	80	0.826	3.3
-1+1-1	40	265	20	0.964	6
+1+1-1	120	265	20	0.973	6.1
00-1	80	245	20	0.979	5.4
0-10	80	225	80	0.842	3.3
000	80	245	80	0.892	3.4

Based on the Table 3-7, it was found that the samples that ranked top 3 (as marked in yellow) in the shear strength among the 16 studied groups happen to be reflowed with the profiles that have top 3 highest reflow cooling down rate (as marked in green), while the samples that ranked bottom 3 (as marked in brown) in terms of the shear strength happen to have undertaken the reflow profiles that have bottom 3 cooling down rate (as marked in light blue). A trend chart was drawn based on the Table 3-7, as shown in Figure 3-21.

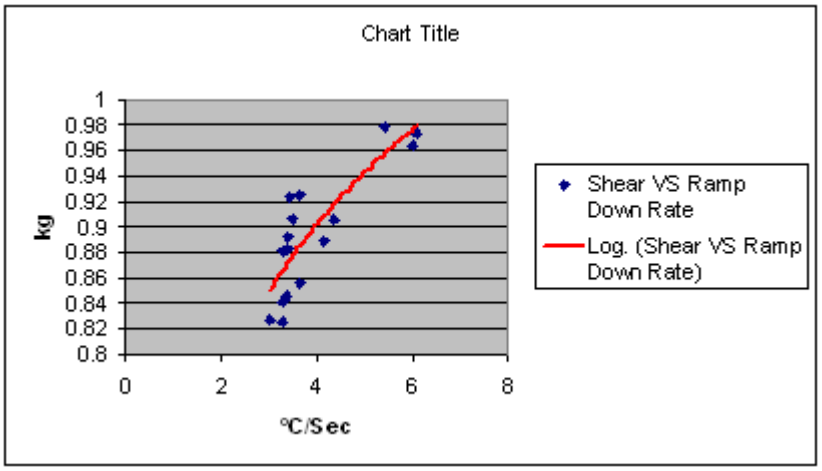


Figure 3-21 3D-WLCSP First Level Assembly Package Shear Strength VS. Reflow Cooling Down Rate

From Figure 3-21, it can be seen that a higher reflow cooling down rate tends to produce higher package shear strength. However, the reflow cooling down rate should not be set as high as possible due to the fact that a too high cooling down rate may produce excessive thermal stress on the package reflowed, which may cause die crack or substrate delamination. A cooling down rate of 2-5°C/s is normally seen in the electronics assembly reflow process.

3.3.4 Reflow Profiles and Flux Amount on Package Yield with Air Reflow Environment

Nitrogen as an inert gas can prevent the solder and the pad surface from re-oxidation during reflow. However, the operation with Nitrogen tends to be much more costly than the operation with air. Thus, for the 3D-WLCSP manufacturing cost perspective, it is necessary to learn if there is any impact of the reflow profiles on the flip chip on silicon assembly yield with the air reflow environment.

Flux amount is also a factor that may affect the flip chip assembly yield while reflowed in air. The flux amount on the solder bumps can be controlled by the flux dip height. In this study, four different reflow profiles were evaluated, with three levels of dip height (15um, 30um and 45um) tested on each profile. The DOE layout as well as the assembly yield under each condition is shown in Table 3-8. Quadrant 3 test vehicles were used in this study.

Table 3-8 First-level Assembly Yield with Air Reflow Environment and Different Flux Dip Height

Reflow Profile	Soak Time (s)	Peak Temp (C)	TAL (s)	Dip Height (um)	Yield
				15	0/8
				30	0/8
-1 -1 -1	40	225	20	45	9/16
				15	0/4
				30	0/4
0 0 0	80	245	80	45	0/6
				15	9/10
				30	9/9
-1 +1 +1	40	265	140	45	7/7
				15	10/10
				30	8/8
0 0 -1	80	245	20	45	8/8

From Table 3-8, it can be seen that the assembly yield is sensitive to the reflow profile condition in the air reflow environment. Samples assembled at different reflow profiles show different yield at the same flux dip height. At the dip flux height of 45 μm, for example, two reflow profiles showed 100% yield, while the other two reflow profiles showed yield loss. It was also found that the yield is sensitive to the flux dip height in certain reflow profiles while insensitive to the flux dip height in other profiles. For example, samples reflowed under condition (0 0 -1) and (0 0 0) had fixed yield (100% in

the former and 0% in the latter condition) regardless of the flux dip height. From the results shown in assembly condition (-1 -1 -1) and (-1 +1 +1), it can be found that a higher flux dip height results in a higher yield.

Efforts have been made to understand the cause of the difference on yield on different reflow profiles at the same flux dip height. At 45μm dip height, the reflow profile condition and yield can be outlined as Table 3-9.

Table 3-9 Reflow Profile Conditions and Yield at 45μm Dip Height

Profile	Yield (45 μm)
0 0 0	0%
−1 −1 −1	56.25%
0 0 −1	100%
−1 +1 +1	100%

Figure 3-22 shows a cross section view of one solder joint reflowed under the profile condition (-1 -1 -1), with 45μm dip height. The solder did not wet on the pad as no intermetallic compound (IMC) was formed in between the solder and pad surface metal. One possible reason is that the flux was not present when the solder started to melt (above >217°C), so that the solder could not wet on pad and form the solder joint.

Figure 3-23 shows the reflow profile of the condition (0 0 -1). The time named “flux burn time” was calculated from the start of the soak zone till the beginning of the solder melting temperature. During this period of time, the flux was activated to remove the oxidation on the solder surface as well as on the pad. In the meanwhile, the flux was burned and evaporated by the heat inside the oven. Therefore, if all the flux is completely

consumed before the package enters the solder melting stage (>217°C), no flux will be present to remove the re-oxidation on the pad and solder. In this case, the yield loss could occur, as the solder does not wet on the pad easily without the help of flux for the oxidation removal during the reflow stage.

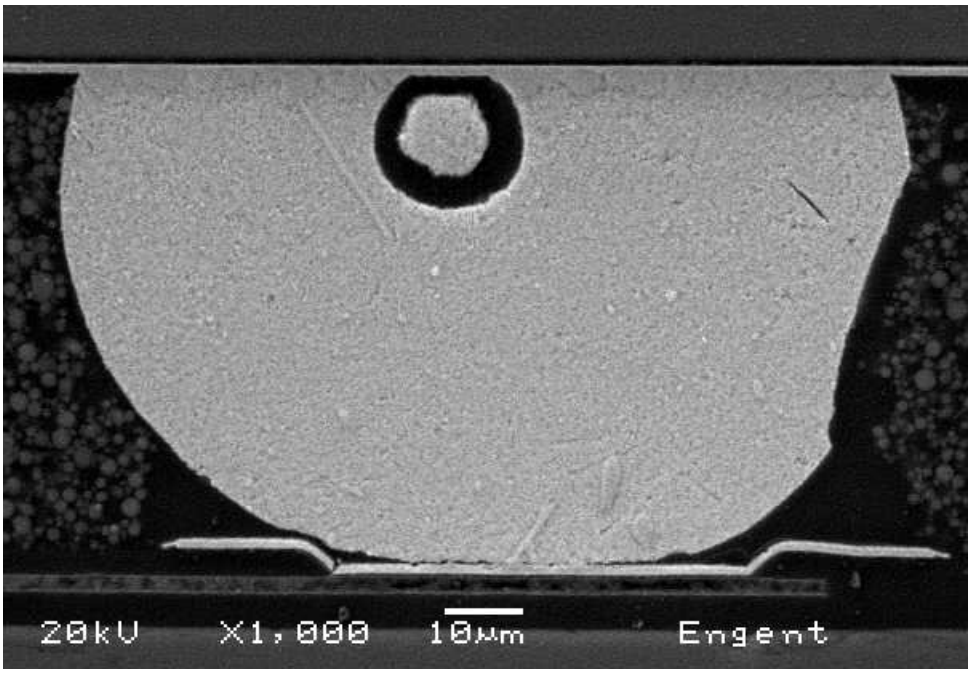


Figure 3-22 Solder Joint Reflowed in Air at Condition (-1 -1 -1) with 45µm Dip Height

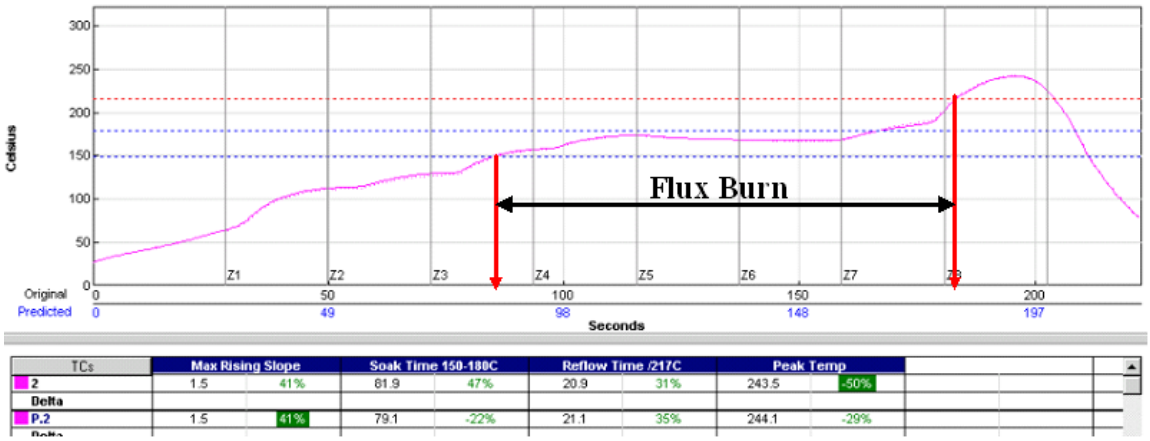


Figure 3-23 Flux Burn Time Calculation

Table 3-10 shows the reflow profile and the corresponding flux burn time as well as the assembly yield with 45µm dip height under the air reflow environment. It was found that the flux burn time is negatively correlated with the package assembly yield, which is shown in Figure 3-24.

Table 3-10 Reflow Profile Conditions, Flux Burn Time and Yield at 45µm Dip Height

Profile	Start of Soak to Start of Solder Melting Temp (sec)	Yield (45um Dip)
0 0 0	162	0%
-1 -1 -1	110	56.25%
0 0 -1	95	100%
-1 +1 +1	75	100%

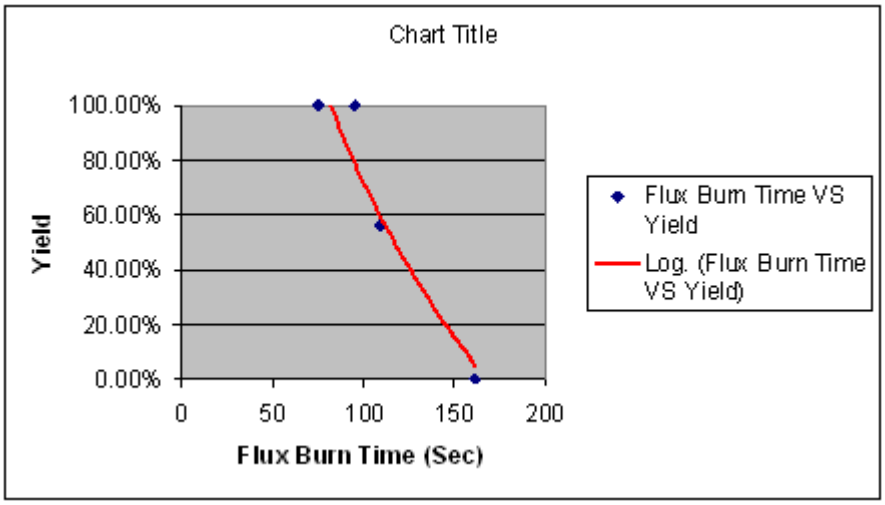


Figure 3-24 Yield VS. Flux Burn Time on Quadrant 3 Package Reflowed in Air

This indicates that in order to achieve high yield in the air reflow environment, a shorter time span from the start of soak zone to the start of solder melting temperature (flux burn time) is desired, so that there will be enough flux material remained to help to

solder wetting activity during the reflow zone. However, the flux burn time should not be as short as possible because an adequate soak time is required before the solder enters the melting stage, to ensure that the whole assembly (especially if the thermal mass is big) is evenly heated and the flux is activated to remove existing oxidation.

3.3.5. Reflow Profiles on Package Reliability Performance

In order to understand the impact of reflow parameters toward flip chip on silicon package thermal reliability performance, 30 Quadrant 1 packages were assembled for each of the fifteen reflow profiles shown in Table 3-3. No underfill was applied after the assembly. All the packages were then subjected to liquid-to-liquid thermal shock (LLTS) testing (-55ºC / +135ºC, 5 minutes dwell time on each side). The electrical continuity across the daisy chain on each package was monitored every 100 cycles. The detailed results up to 2500 cycles are shown in Table 3-11. Only one sample from the assembly condition (0 –1 0) failed at the end of 2200 cycles, based on the electrical continuity measurement. The reliability test result shows that the silicon on a silicon flip chip assembled package is very robust, as no CTE mismatch between die and substrate is present in this structure.

Figure 3-25 shows the cross sectional view of a solder joint that survived 2500 cycles of LLTS testing. Figure 3-26 shows the solder joint that failed the LLTS testing. Although there is no CTE mismatch between the silicon die and silicon substrate, as no underfill was applied, the testing liquid may be trapped inside the gap between die and substrate. With rapid temperature changes, the testing liquid and solder may have a CTE

mismatch along the Z-axis. This can explain the failure that was captured by the cross-section as shown in Figure 3-16.

Table 3-11 Quadrant 1 LLTS Results Vs. Reflow Profile Condition

Pattern	Before Test	500 cycles	1000 Cycles	1500 Cycles	2000 Cycles	2500 Cycles
+1-1-1	0/30	0/30	0/30	0/30	0/30	0/30
+1+1+1	0/30	0/30	0/30	0/30	0/30	0/30
+1 0 0	0/30	0/30	0/30	0/30	0/30	0/30
0 0 0	0/30	0/30	0/30	0/30	0/30	0/30
-1 0 0	0/30	0/30	0/30	0/30	0/30	0/30
0 0-1	0/30	0/30	0/30	0/30	0/30	0/30
0 0+1	0/30	0/30	0/30	0/30	0/30	0/30
+1-1+1	0/30	0/30	0/30	0/30	0/30	0/30
-1+1-1	0/30	0/30	0/30	0/30	0/30	0/30
-1-1-1	0/30	0/30	0/30	0/30	0/30	0/30
0+1 0	0/30	0/30	0/30	0/30	0/30	0/30
+1+1-1	0/30	0/30	0/30	0/30	0/30	0/30
-1+1+1	0/30	0/30	0/30	0/30	0/30	0/30
0 0 0	0/30	0/30	0/30	0/30	0/30	0/30
-1-1+1	0/30	0/30	0/30	0/30	0/30	0/30
0 -1 0	0/30	0/30	0/30	0/30	0/30	1/30

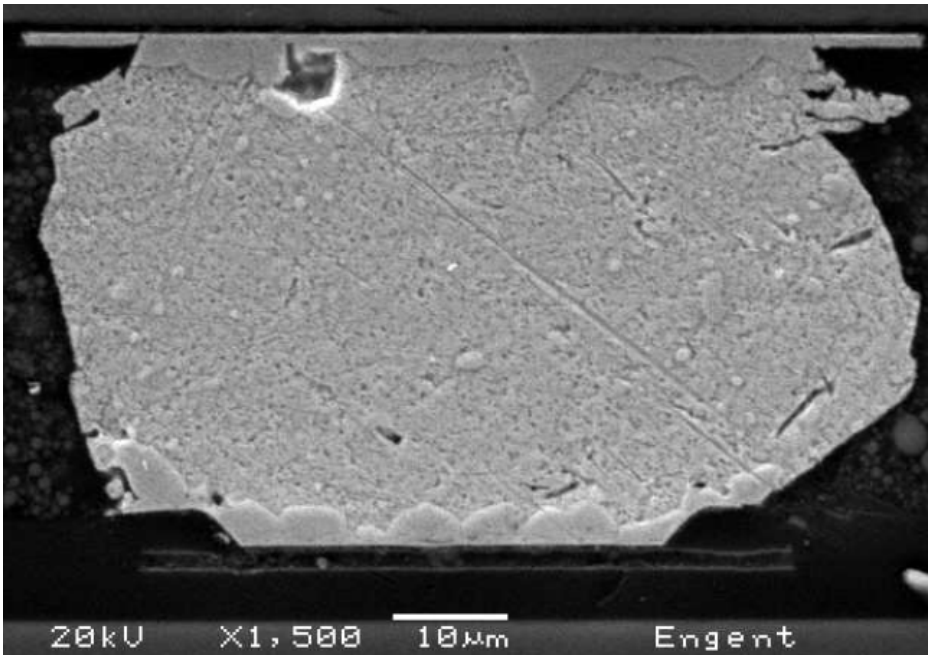


Figure 3-25 Good Solder Joint after LLTS testing

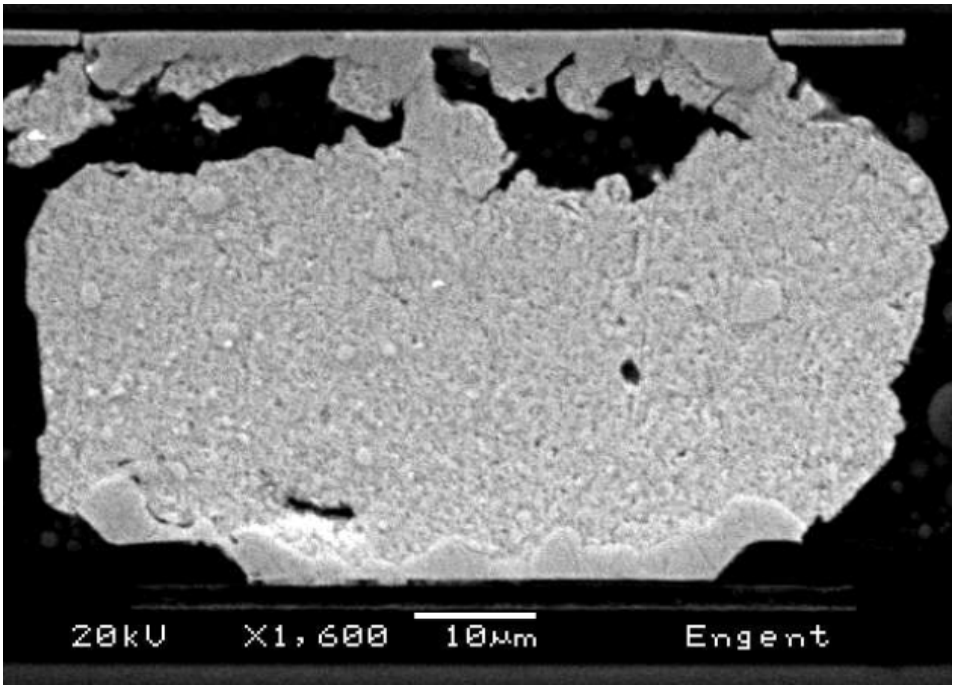


Figure 3-26 Failed Solder Joint after LLTS testing

3.4. Summary

This chapter discussed the research on the impact of reflow profile parameters toward fine-pitch flip chip on silicon assembly. Three reflow parameters including soak time, peak temperature and time above liquidus were investigated. The Response Surface Methodology was used to design the experiment, in order to investigate the quadratic effect of each profile parameter and the interactions among factors.

It was found that 100% yield could be achieved in a very wide reflow processing window, if the package is reflowed in a Nitrogen environment. However, the flip chip assembly yield is sensitive to the reflow profile parameters if the package is reflowed in the air environment. The higher “flux burn” time leads to lower package yield when packages are reflowed in the air environment.

The intermetallic compound was measured with Simpson's rule on samples reflowed under each of the fifteen reflow profiles. It was found that the time above liquidus is the only significant factor at 95% level among the three factors in the studied reflow matrix on the IMC thickness.

None of the three reflow profile parameters investigated was found to be significant in terms of the package shear strength. However, it was found that the reflow profile cooling down rate has a positive co-relationship with the package shear strength. The higher cooling down rate tends to yield higher package shear strength.

LLTS reliability testing is conducted on samples assembled with fifteen different reflow profiles following the RSM DOE design. Only one sample failure was observed after 2500 cycles, which indicates a very robust thermal reliability performance for the 3D-WLCSP first level silicon-on-silicon packages.

CHAPTER 4 3D-WLCSP FIRST-LEVEL CAPILLARY UNDERFILL FLIP CHIP ASSEMBLY PROCESS DEVELOPMENT

4.1 Introduction

The first-level assembly of the 3D-WLCSP technology is realized by face-to-face bonding of the flip chip die to the CSP substrate on the wafer format. Although the reliability performance of the first-level assembly without underfill is satisfactory as discussed in Chapter 3, underfill material is still required for the first-level assembly due to the concern that mechanical stress may be induced on flip chip solder joint during dicing activities on the assembled wafer, in order to singulate the first level assembled packages from the wafer. Furthermore, during second-level assembly, the first-level assembled package will endure the mechanical stress by pick and placement activities. Other mechanical impact induced from transportation or drop events also require that the packages be underfilled to overcome the stress and strain applied to the solder joints during such events.

As discussed in Chapter 1, the flip chip assembly can be grouped into two main categories in terms of the underfill application. The first method is the capillary flow underfill process. For the 3D-WLCSP at the first-level, the capillary flow underfill assembly process can be demonstrated in Figure 4-1.

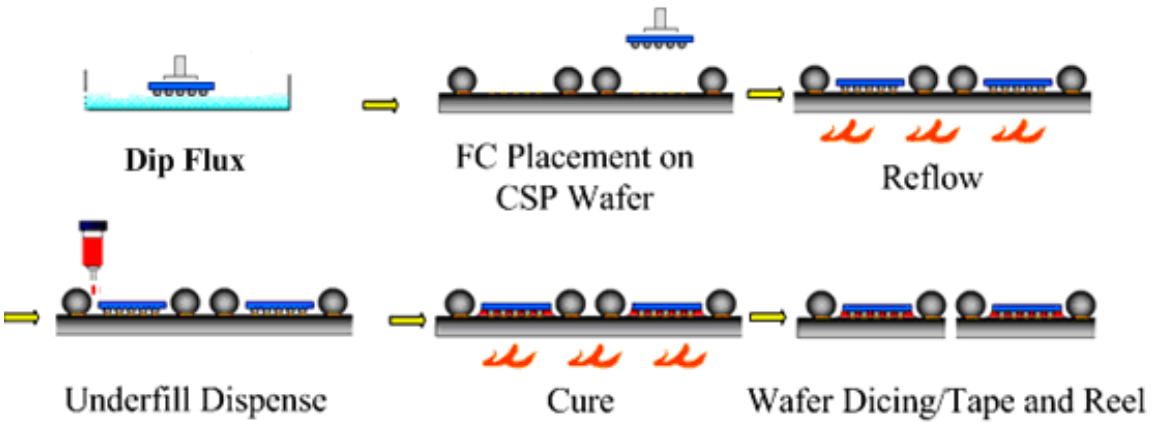


Figure 4-1 3D-WLCSP First-level Capillary Underfill Assembly Process

No-clean tacky flux was used in the 3D-WLCSP first-level assembly, in order to improve the process throughput. Although no-clean flux does not leave the corrosive residues as most of the clean-required flux materials do, the no-clean flux may still leave residues that can cover the solder joints and cause underfill voids as well as delamination. Zhang, Chan, and Chiang reported [47] that the Tg value of the underfill decreased with the presence of residues from the no-clean flux that they studied. The CTE above Tg in the underfill was found to increase significantly. All these changes are expected to affect the package long-term reliability performance. Thus, a goal of this research is to test different flux/underfill materials and identify the best performing flux/underfill combination to enhance the 3D-WLCSP first level package long-term reliability performance.

4.2 Experimental Approach

To accomplish the goal of selecting the most compatible and best performing flux/underfill material combination, a full-factorial design of experiment (DOE) was conducted. Three factors, flux, underfill and package layouts, were investigated. Three flux materials and four underfill materials were selected to be tested in this study, based on the manufacturing recommendations and previous experience with Pb-free applications by Engent Inc. The three tacky flux materials used in this experiment were designed for the dip fluxing process for lead free flip chip on board (FCOB) applications. [48,49]. The underfill materials' properties are listed in Table 4-1.

Table 4-1 Underfill Information for the Flux/Underfill Compatibility Study

Underfill	CTE (ppm/°C)
1	50
2	210
3	22
4	28

Test vehicle Quadrant 2 and Quadrant 4 were used in the study conducted in this chapter. As shown in Table 3-1, the Quadrant 2 has a bump pitch of 85μm while Quadrant 4 has a bump pitch of 200μm. The bumps on Quadrant 2 are also smaller than those on Quadrant 4. The Quadrant 2 and Quadrant 4 test vehicles are used to compare

the impact of design layout toward the package yield and long-term reliability performance.

A 2x3x4 full-factorial DOE with 24-factorial level combinations was outlined as shown in Table 4-2.

Table 4-2 Flux/Underfill Compatibility Study randomized DOE Layout

Quadrant	Flux	Underfill
2	B	3
2	B	4
1	C	4
2	A	4
1	B	4
1	C	2
2	B	2
2	C	3
2	A	2
1	C	1
1	A	2
1	A	1
1	B	3
1	A	4
1	A	3
1	B	2
1	C	3
2	A	3
2	C	2
2	C	1
2	B	1
2	A	1
2	C	4
1	B	1

A standard dip flux, capillary underfill process was adopted in the assembly process. The ESEC Micron 2 chip placement machine was used to pick and place the flip chip component. The reflow was performed in a nitrogen environment in a Rehm RN38 convection reflow oven. The reflow profile used for lead-free flip chip assembly in this study is shown in Figure 4-2.

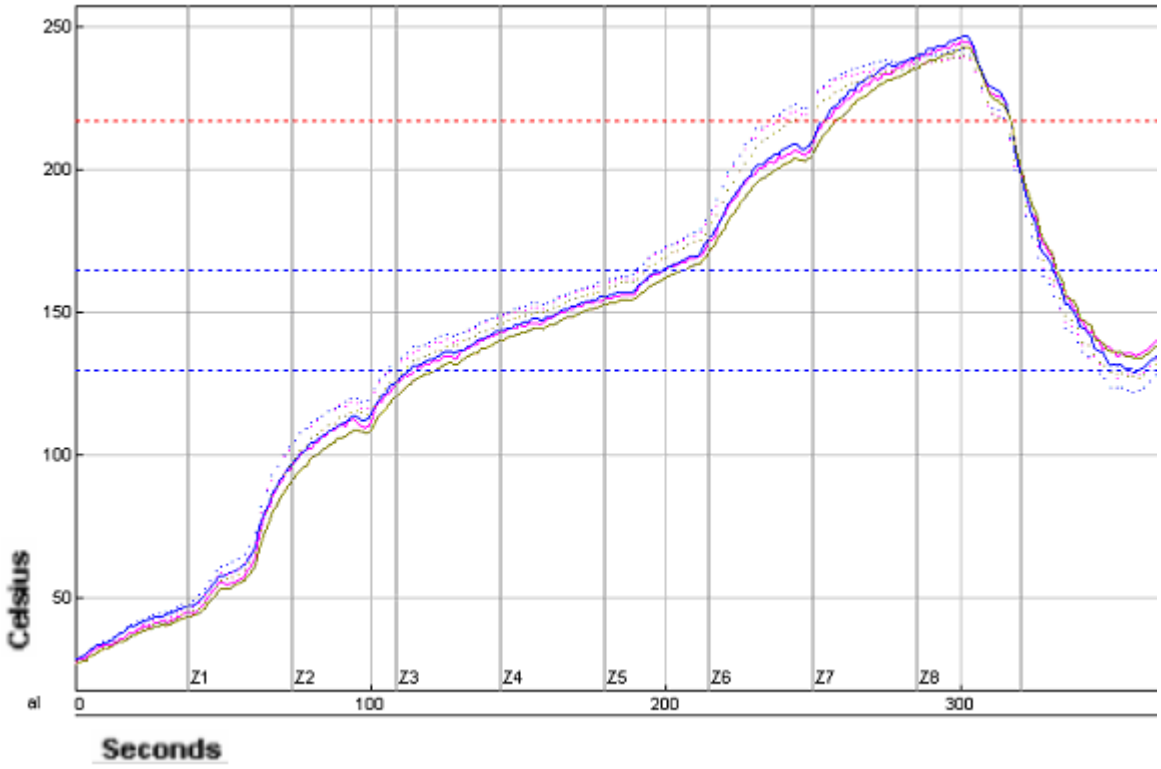


Figure 4-2 Example of 3D WLCSP First-level Assembly Reflow Profile

Two capillary underfill dispense methods were used in this experiment. The two methods are a needle dispensing method and a jet dispensing method. The needle dispensing method is a traditional underfill dispensing method that utilizes the pump-driven needle. The fluid is in contact with both needle and the target location during dispensing. By contrast the jet dispensing method is a non-contact method since the underfill droplets leave the dispense nozzle before contacting the target substrate. Underfills 1 through 3 were dispensed using the needling dispensing method with positive displacement pump and a 25 gage needle on a Camalot 5700 dispense machine. Underfill 4 was dispensed using a jetting technique on an Asymtek DJ-9000 jet dispenser.

Figure 4-3 shows the optical image of Quadrant 4 ACT on Quadrant 4 PA assembly with a needle-dispensed underfill. Figure 4-4 shows the optical image of a Quadrant 2 assembly with jet-dispensed underfill. By comparison of Figure 4-3 and 4-4, it can been seen that the jetting technique has better underfill volume control over the standard needle dispensing method in that the wetted area around the die with the jetting dispense method was smaller than that with the needle dispense method. This volume control is especially important considering that the excessive amount of underfill may encroach on the solder balls of the WLCSP substrate, which would cause a solder wetting problem for the second-level assembly.

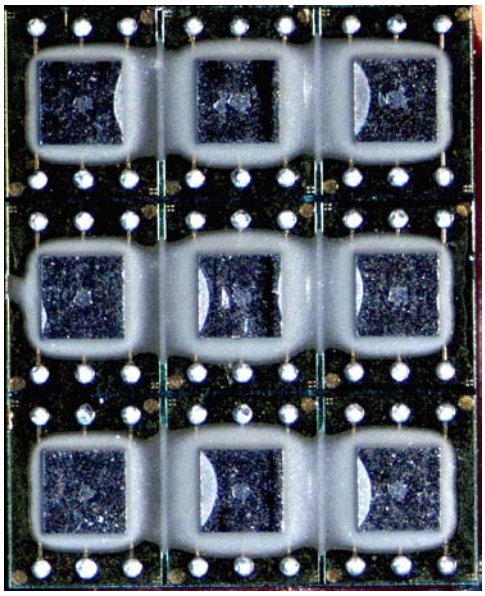


Figure 4-3 WLCSP Assembled with Underfill 3 Dispensed via Positive Displacement Pump

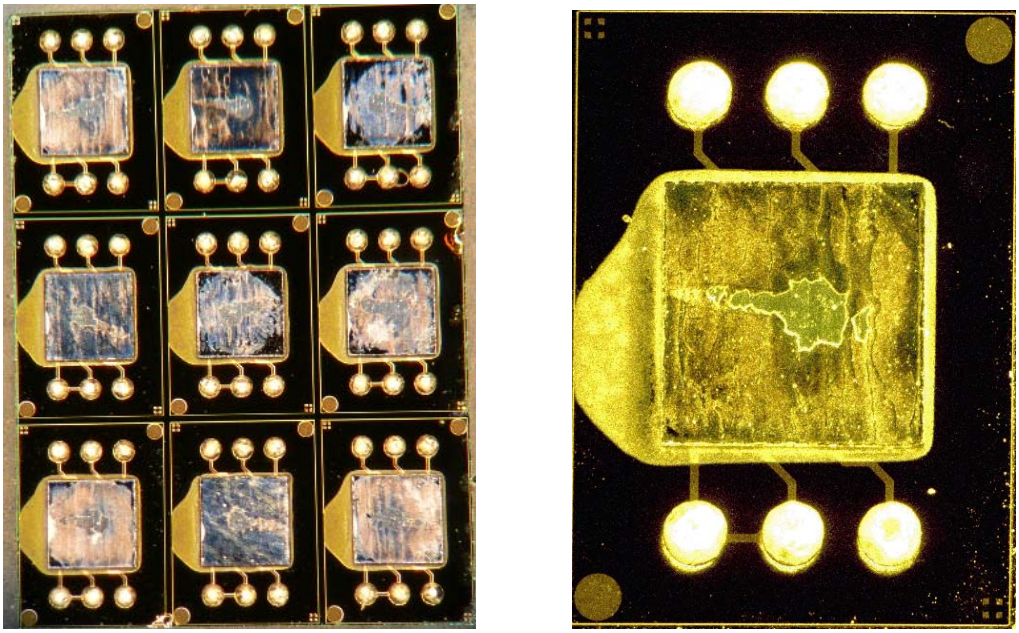


Figure 4-4 Examples of Underfill 4 Dispensed via Jet Deposition

A few cross sections after assembly were performed to check the condition of solder joints as well as the underfill cure condition. Figure 4-5 shows a cross section image of a Quadrant 2 bump taken by Scanning Electron Microscope (SEM) showing a high quality flip chip interconnect after reflow. As shown in the image, the stand-off height of the solder bump after collapse is only 28 microns. Such a small gap requires the underfill to have very small particle sizes to fill the gap between the silicon die and the silicon substrate. Figure 4-6 shows the overall view of a Quadrant 2 first-level package cross section after reflow.

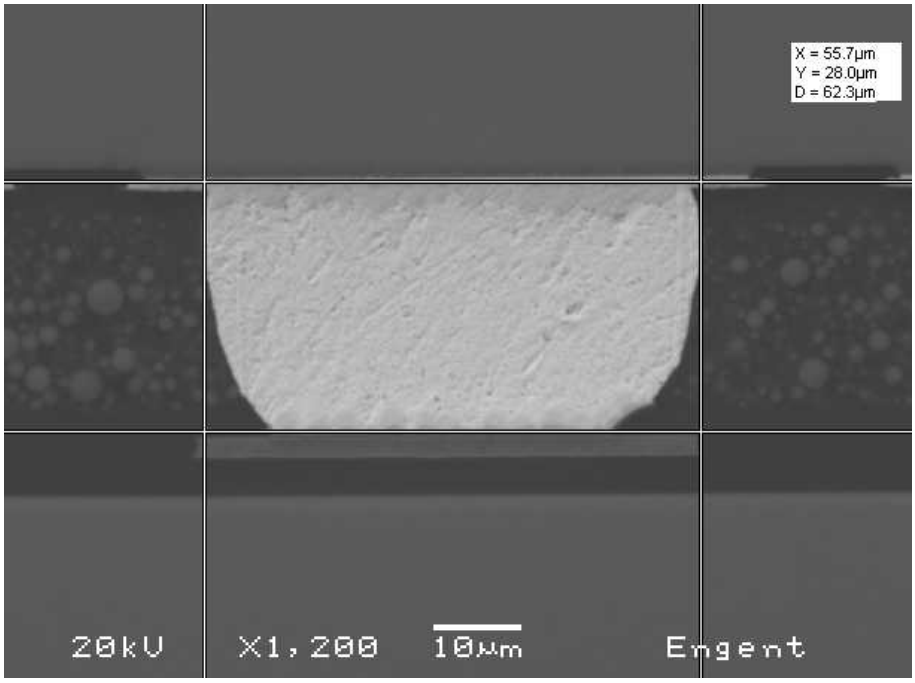


Figure 4-5 SEM Image of a Quadrant 2 Bump after Reflow

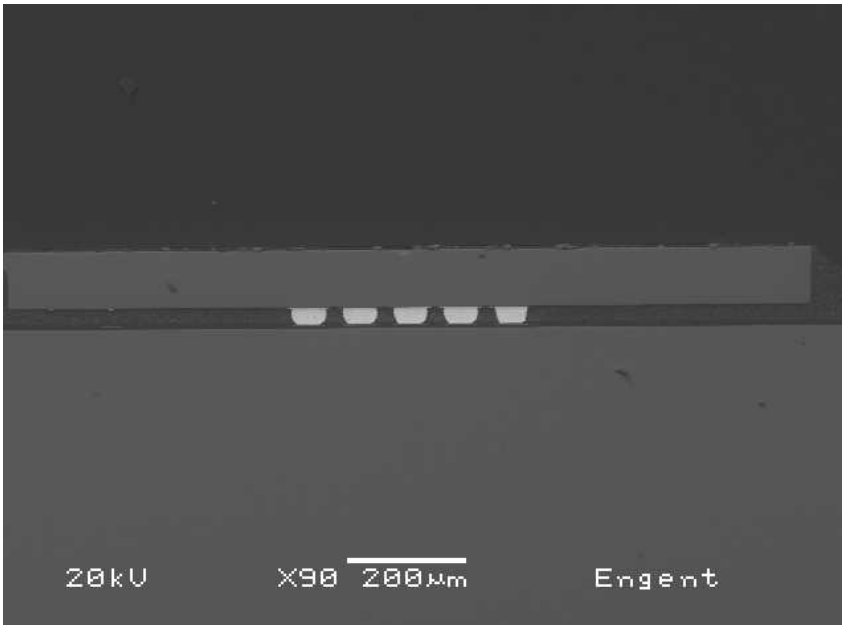


Figure 4-6 Cross Section of a Quadrant 2 First-level Assembly

After initial electrical resistance measurement, the packages assembled were split into two batches for different tests. The first batch of assembly was subjected to

Liquid-to-Liquid Thermal Shock (LLTS) testing temperature settings of -40ºC and 125ºC for 1000 cycles and the other batch underwent autoclave testing with settings of 121ºC, 100% RH, 2 atm pressure for 96 hours. Those tests were performed as a screening test to select the best underfill/flux combination for the first-level assembly of the proposed 3D-WLCSP structure. The detailed LLTS testing matrix with the sample tested for each factorial level combination is shown in Table 4-3.

Table 4-3 LLTS Test Matrix

	Q2 Assembled Packages	Q4 Assembled Packaged
Flux A Underfill 1	13	18
Flux A Underfill 2	8	27
Flux A Underfill 3	15	27
Flux A Underfill 4	16	18
Flux B Underfill 1	18	18
Flux B Underfill 2	18	27
Flux B Underfill 3	17	27
Flux B Underfill 4	24	27
Flux C Underfill 1	17	18
Flux C Underfill 2	15	18
Flux C Underfill 3	16	18
Flux C Underfill 4	16	25

The electrical resistance across the daisy chain on each component was measured every 200 cycles, with +/- 10% change in the resistance as pass/fail criteria. The C-mode scanning acoustic microcopy (C-SAM) images were also taken every 200 cycles during cycling, to identify if voiding or delamination occurs in the underfill during the thermal shock testing.

4.3 Experimental Results

4.3.1 3D-WLCSP First Level Assembly Flux/Underfill Compatibility Testing

Table 4-4 shows the 1000 cycles of LLTS reliability test-data for Quadrant 2 assembly and table 4-5 shows the 1000 cycles of LLTS reliability test-data for Quadrant 4 assembly.

Table 4-4 Quadrant 2 LLTS Reliability Data

	Cycles				
Material Combination	**200**	**400**	**600**	**800**	**1000**
Flux A Underfill 1	**0/13**	**2/13**	**6/13**	**11/13**	**13/13**
Flux A Underfill 2	**8/8**	**8/8**	**8/8**	**8/8**	**8/8**
Flux A Underfill 3	**0/15**	**1/15**	**3/15**	**5/15**	**6/15**
Flux A Underfill 4	**0/16**	**0/16**	**0/16**	**0/16**	**1/16**
Flux B Underfill 1	**0/18**	**3/18**	**9/18**	**15/18**	**15/18**
Flux B Underfill 2	**18/18**	**18/18**	**18/18**	**18/18**	**18/18**
Flux B Underfill 3	**0/17**	**0/17**	**0/17**	**0/17**	**1/17**
Flux B Underfill 4	**0/24**	**1/24**	**1/24**	**1/24**	**3/24**
Flux C Underfill 1	**0/17**	**8/17**	**13/17**	**17/17**	**17/17**
Flux C Underfill 2	**24/24**	**24/24**	**24/24**	**24/24**	**24/24**
Flux C Underfill 3	**0/16**	**1/16**	**1/16**	**1/16**	**1/16**
Flux C Underfill 4	**0/16**	**0/16**	**0/16**	**0/16**	**3/16**

Table 4-5 Quadrant 4 LLTS Reliability Data

	Cycles				
Material Combination	**200**	**400**	**600**	**800**	**1000**
Flux A Underfill 1	**0/18**	**0/18**	**0/18**	**1/18**	**2/18**
Flux A Underfill 2	**27/27**	**27/27**	**27/27**	**27/27**	**27/27**
Flux A Underfill 3	**1/27**	**1/27**	**2/27**	**4/27**	**6/27**
Flux A Underfill 4	**0/18**	**0/18**	**0/18**	**0/18**	**0/18**
Flux B Underfill 1	**2/18**	**5/18**	**9/18**	**13/18**	**14/18**
Flux B Underfill 2	**27/27**	**27/27**	**27/27**	**27/27**	**27/27**
Flux B Underfill 3	**0/27**	**0/27**	**0/27**	**0/27**	**2/27**
Flux B Underfill 4	**0/27**	**0/27**	**0/27**	**0/27**	**0/27**
Flux C Underfill 1	**0/18**	**1/18**	**1/18**	**4/18**	**5/18**
Flux C Underfill 2	**18/18**	**18/18**	**18/18**	**18/18**	**18/18**
Flux C Underfill 3	**0/18**	**0/18**	**0/18**	**0/18**	**0/18**
Flux C Underfill 4	**0/25**	**1/25**	**1/25**	**1/25**	**1/25**

Figure 4-7 shows the Weibull plot for Quadrant 2 assemblies with underfill 3 and underfill 4. Figure 4-8 shows the LLTS Weibull plot for Quadrant 4 assemblies with underfills 3 and 4.

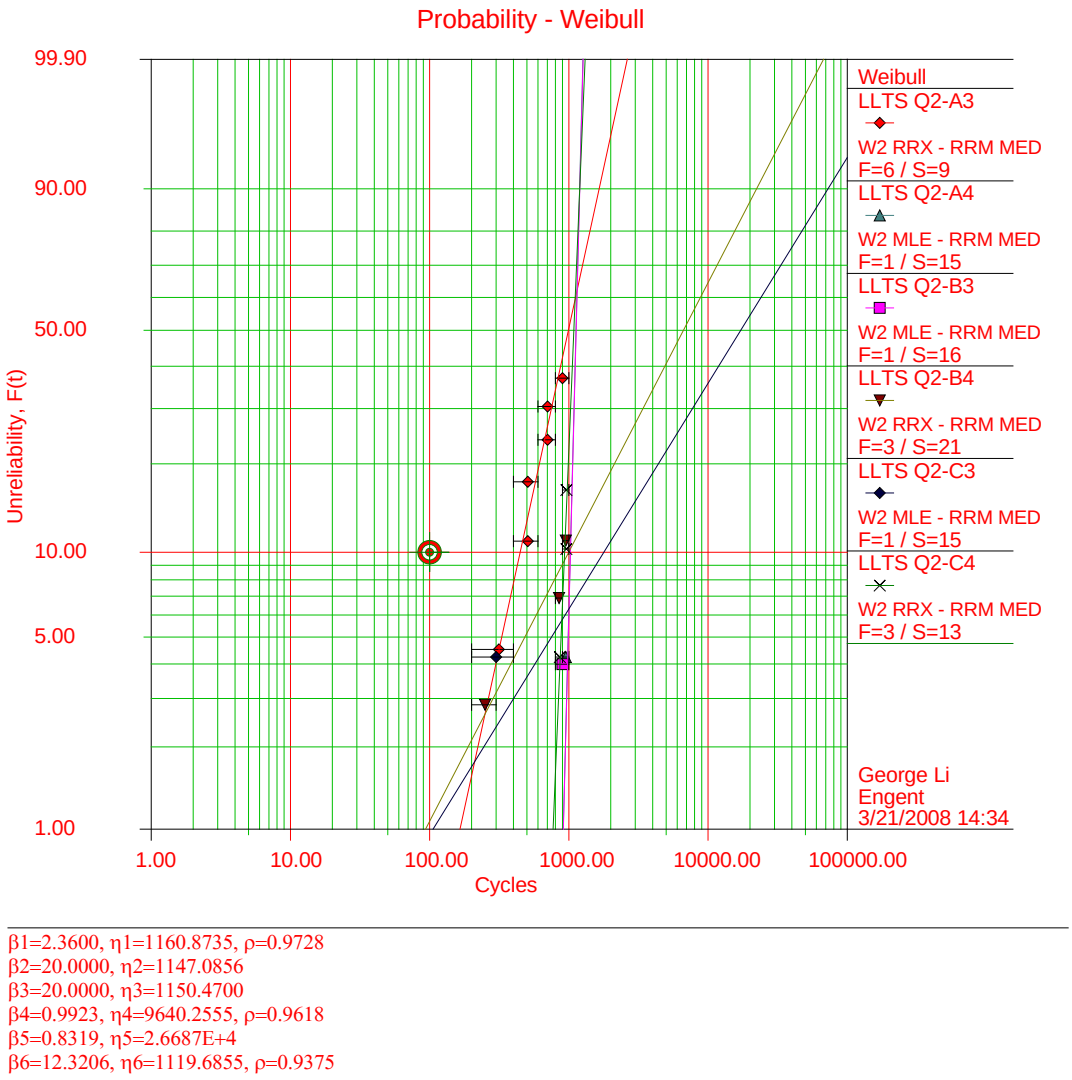


Figure 4-7 Weibull Plot for Underfill 3 and Underfill 4 with Flux A, B, and C for Quadrant 2 LLTS Testing

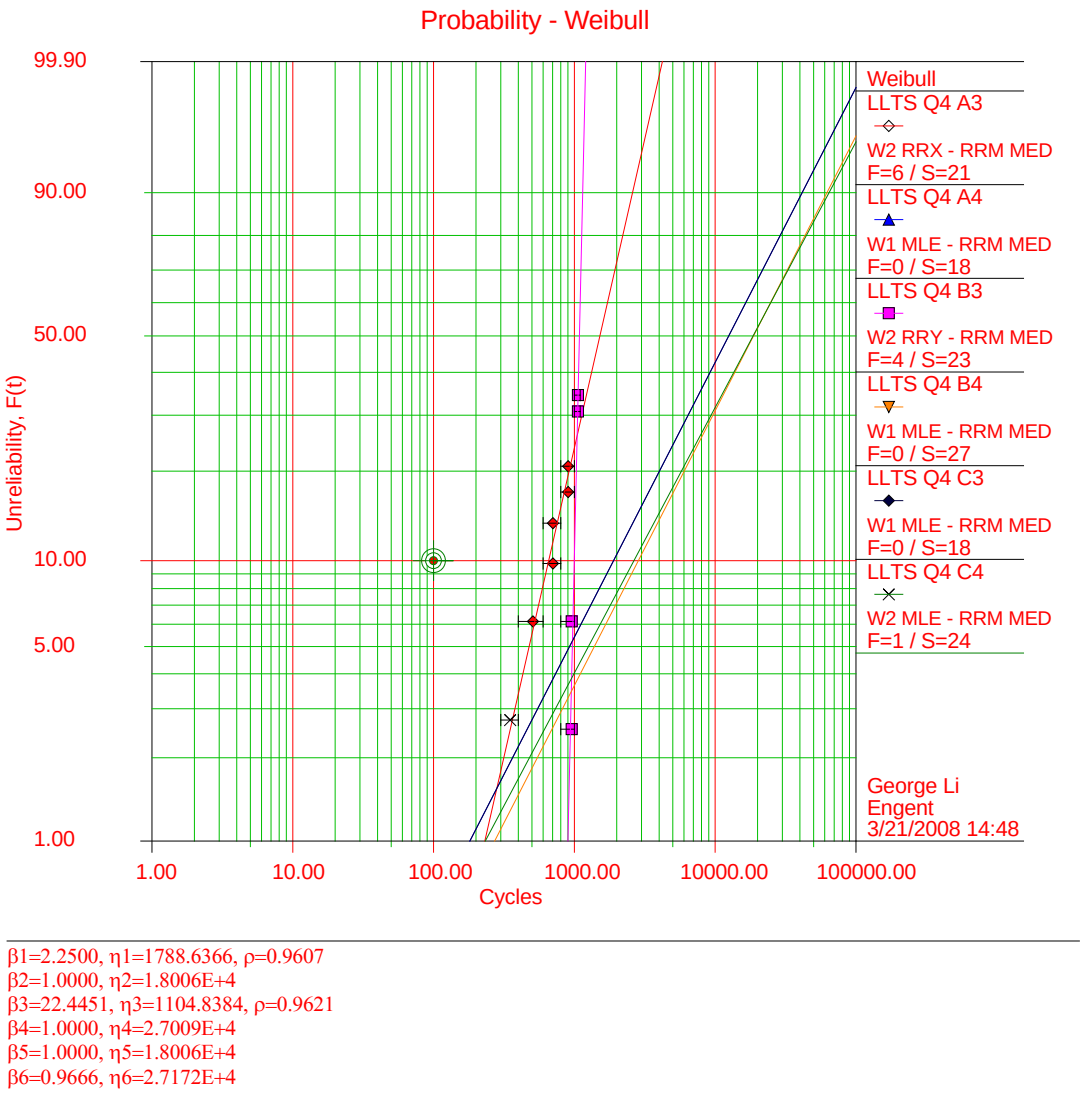


Figure 4-8 Weibull Plot for Underfills 3 and 4 with Fluxes A, B, and C for Quadrant 4 LLTS Testing

Based on the 1000 cycles reliability data (percentage of units passed 1000 cycles) shown in table 4-4 and 4-5, an analysis of variance (ANOVA) was performed, to identify the impact of each factor toward package reliability performance. Figure 4-9 shows the results of the ANOVA analysis by Minitab software.

General Linear Model: Reliability versus Underfill, Flux, Quadrant

Factor	Type	Levels	Values
Underfill	fixed	4	1, 2, 3, 4
Flux	fixed	3	1, 2, 3
Quadrant	fixed	2	1, 2

Analysis of Variance for Reliability, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Underfill	3	3.54964	3.54964	1.18321	50.35	0.000
Flux	2	0.00617	0.00617	0.00308	0.13	0.879
Quadrant	1	0.20657	0.20657	0.20657	8.79	0.025
Underfill*Flux	6	0.15852	0.15852	0.02642	1.12	0.445
Underfill*Quadrant	3	0.28349	0.28349	0.09450	4.02	0.069
Flux*Quadrant	2	0.06485	0.06485	0.03242	1.38	0.321
Error	6	0.14100	0.14100	0.02350		
Total	23	4.41026				

S = 0.153298 R-Sq = 96.80% R-Sq(adj) = 87.74%

Figure 4-9 Analysis of Variance (ANOVA) on the Flux/Underfill Compatibility LLTS Testing Data

ANOVA results indicate that underfill is the most significant factor towards the first level package reliability performance, as the P-value of underfill is zero. The flux is not significant towards reliability performance. The flip chip package layout design, which is represented by the quadrant number, is also significant at 95% confidence level with a P-value of 0.025. As for the interactions among factors, only the interaction between underfill and quadrant is significant at a confidence level of 93%, with a P-value of 0.069. This means that the interaction between underfill and quadrant also plays a role in the package long-term reliability.

Underfill 3 and 4 are more typical flip chip in package materials, which fit the first-level interconnection application. underfill 1 and 2 are more typically used in underfilling CSP devices. The reason that underfilll 1 and 2 were evaluated was for their lower stress profiles, lower filler content and faster curing profiles, which has the potential to increase the throughput. underfills 1 and 2 cured in 5 to 10 minutes while underfills 3 and 4 required 30 minutes to two hours to cure. It was found from the reliability testing data that the flux/underfill combinations of A4, B4 and C3 show best reliability performance among all the combinations, both for Quadrant 2 and Quadrant 4. Other combination like B3 also show good reliability with only 1 failure at the end of 1000 LLTS cycles.

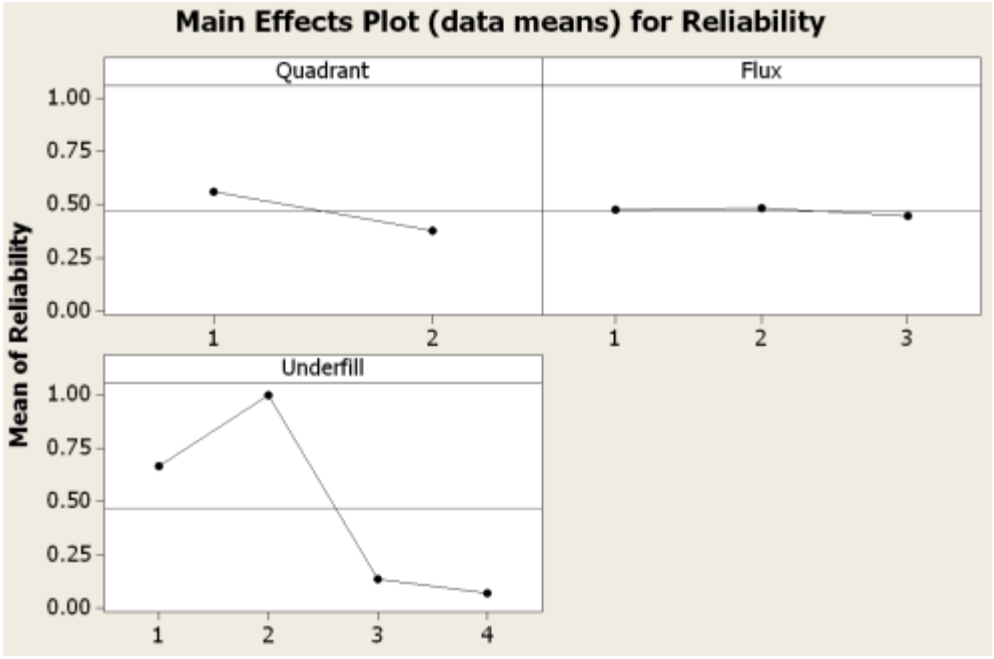


Figure 4-10 Main Effect Plot for Flux/Underfill Compatibility Study

A main effect plot drawn by Minitab software is shown in Figure 4-10. It can be seen that underfill 3 and 4 outperformed underfill 1 and 2 in reliability results. The underfill 3 and underfill 4 tended to have higher modulus and lower CTE relative to the underfill 1 and underfill 2. The underfill 3 and underfill 4 also have closer CTE value to the CTE value of Pb-free solder material (around 24 ppm/°C) than do the underfill 1 and underfill 2, as shown in Table 1. It is obvious that underfill 3 and underfill 4, which have a closer CTE value to the SAC Pb-free solder than underfill 1 and underfill 2, shall better absorb the stress on solder joints induced by the CTE mismatch between Pb-free solder and silicon during temperature cycling, as these underfills tend to have a similar degree of expansion and contraction as do the Pb-free solder material. This can also explain the reason why underfill 3 and underfill 4 have much better reliability performance than do underfill 1 and underfill 2.

Cross sections on sampling parts were performed to understand the root cause for the failed parts. Figure 4-11 shows the cross section of a failed Quadrant 4 bump after 1000 LLTS cycles with Flux A/Underfill 2. The solder joint fatigue during cycling is the root cause as the crack happened across the solder bulk. This cross section together with others confirmed that the solder joint fatigue during cycling is the main failure mode for the 3D-WLCSP first-level LLTS tested packages.

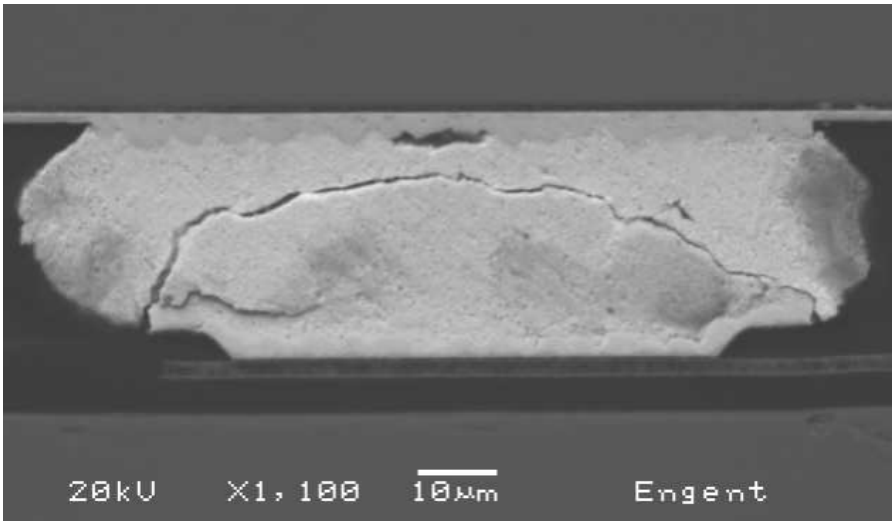


Figure 4-11 Cross section of failed Quadrant 4 bump after LLTS testing assembled with Flux A and Underfill 2

Figure 4-12 shows C-SAM micrograph of a Quadrant 2 Flux B/Underfill 3 assembly, at 0 cycles and after 1000 cycles, respectively. No delamination was observed in any of the underfill 3 or 4 material combinations during the LLTS cycling or autoclave testing.

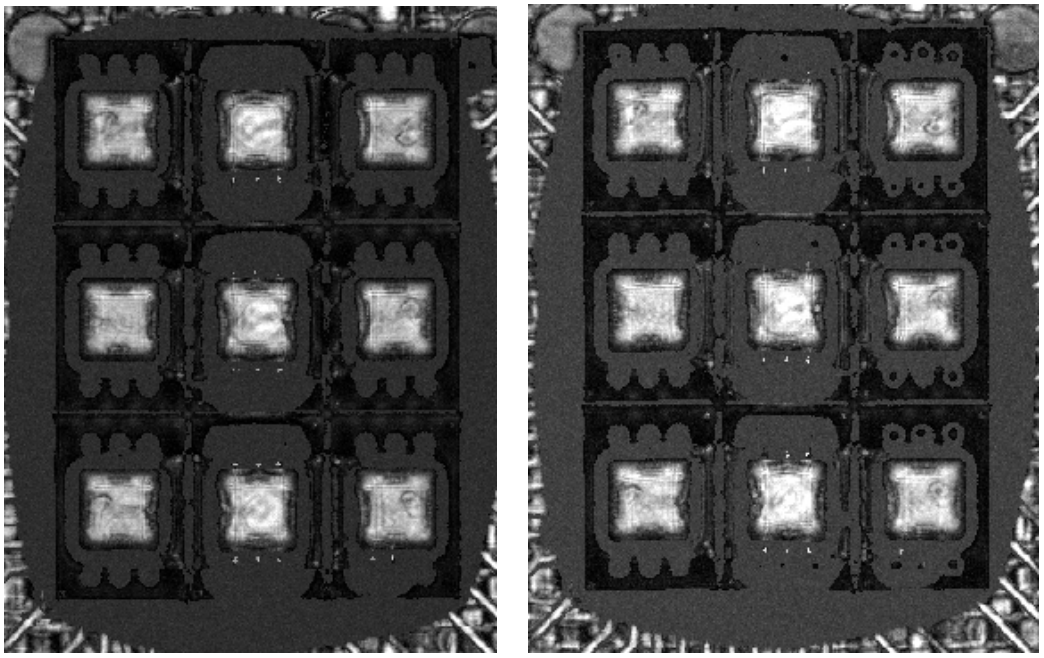


Figure 4-12 C-SAM of Quadrant 2 Flux B/Underfill 3 at 0 cycle (left) and after 1000 cycles (right)

The reliability data of the Autoclave test for Quadrant 2 and Quadrant 4 are shown as Tables 4-6 and 4-7. No failure was observed in either quadrant assemblies during 96 hours of Autoclave testing, using the electrical resistance continuity measurement.

Table 4-6 Quadrant 2 Autoclave Reliability Data

	Aging Hours				
Material Combination	0	24	48	72	96
Flux A Underfill 1	0/11	0/11	0/11	0/11	0/11
Flux A Underfill 2	0/20	0/20	0/20	0/20	0/20
Flux A Underfill 3	0/20	0/20	0/20	0/20	0/20
Flux A Underfill 4	0/9	0/9	0/9	0/9	0/9
Flux B Underfill 1	0/15	0/15	0/15	0/15	0/15
Flux B Underfill 2	0/16	0/16	0/16	0/16	0/16
Flux B Underfill 3	0/18	0/18	0/18	0/18	0/18
Flux B Underfill 4	0/16	0/16	0/16	0/16	0/16
Flux C Underfill 1	0/8	0/8	0/8	0/8	0/8
Flux C Underfill 2	0/17	0/17	0/17	0/17	0/17
Flux C Underfill 3	0/18	0/18	0/18	0/18	0/18
Flux C Underfill 4	0/15	0/15	0/15	0/15	0/15

Table 4-7 Quadrant 4 Autoclave Reliability Data

Material Combination	Aging Hours				
	0	24	48	72	96
Flux A Underfill 1	0/18	0/18	0/18	0/18	0/18
Flux A Underfill 2	0/18	0/18	0/18	0/18	0/18
Flux A Underfill 3	0/18	0/18	0/18	0/18	0/18
Flux A Underfill 4	0/18	0/18	0/18	0/18	0/18
Flux B Underfill 1	0/18	0/18	0/18	0/18	0/18
Flux B Underfill 2	0/18	0/18	0/18	0/18	0/18
Flux B Underfill 3	0/18	0/18	0/18	0/18	0/18
Flux B Underfill 4	0/18	0/18	0/18	0/18	0/18
Flux C Underfill 1	0/18	0/18	0/18	0/18	0/18
Flux C Underfill 2	0/26	0/26	0/26	0/26	0/26
Flux C Underfill 3	0/27	0/27	0/27	0/27	0/27
Flux C Underfill 4	0/16	0/16	0/16	0/16	0/16

4.3.2 3D-WLCSP First Level Assembly Qualification Testing

The LLTS and Autoclave tests served as a screening method to select the best flux/underfill materials. Based on these results, a more comprehensive thermal cycling reliability evaluation of the 3D-WLCSP packaged components at the first-level assembly was carried out in this research. The combination of underfill/flux materials (Flux B/ Underfill 3 and Flux B/Underfill 4) identified in pervious research was used in this study. 28 Quadrant 2 first-level packages and 89 Quadrant 3 first-level packages were assembled, with the underfill 3 applied by a positive displacement pump on the Camalot 5700 dispenser machine with a 25-gauge needle. All the assemblies were then subjected to 5000 cycles air-to-air thermal cycling (AATC) testing (–55 ºC to 125 ºC with a dwell time of 10 minutes on each side per cycle). The electrical continuity of the daisy chain on each component was measured and recorded every 100 cycles, with a 20% change in resistance as the fail/pass criteria. C-mode scanning acoustic microcopy (C-SAM) images were taken every 500 cycles during the AATC testing to check if there is any

delamination in the underfill. The reliability data up to 5000 cycles is shown in Table 4-8 (Quadrant 2 samples were pulled out after 3000 cycles as a majority of Quadrant 2 samples failed by 3000 cycles). Only one Quadrant 3 sample was identified by electrical continuity probing as having failed at the end of 4000 cycles.

Table 4-8 First-level Assembly Qualification Testing Data (Underfill 3 Needle Dispensed)

	Cycles									
Quadrant	500	1000	1500	2000	2500	3000	3500	4000	4500	5000
Q2	0/28	1/28	7/28	12/28	15/28	21/28	N/A	N/A	N/A	N/A
Q3	0/89	0/89	0/89	0/89	0/89	0/89	0/89	1/89	1/89	1/89

Another batch of 99 Q3 first-level packages was assembled with the underfill 4 applied by an Asymtek Jet dispenser utilized, for its better control of the underfill volume. Those assemblies were then subjected to AATC testing at the same profile (–55 ºC to 125 ºC, 10 minutes dwell time on each side, 10 minutes transition time). The reliability data of this batch of assemblies up to 4000 cycles is reported in Table 4-9. No failure was observed throughout the AATC testing by the electrical continuity measurement.

Table 4-9 First-level Assembly Qualification Testing Data (Underfill 4 Jet Dispensed)

	Cycles							
Quadrant	500	1000	1500	2000	2500	3000	3500	4000
Q3	0/99	0/99	0/99	0/99	0/99	0/99	0/99	0/99

Based on the reliability test data in Table 4-8 and Table 4-9, it can be seen that Quadrant 3, which has a larger pitch and bump diameter than those of Quadrant 2, shows robust reliability performance. Quadrant 2 samples had the first failure by the end of 1000 cycles, which also indicates a good package reliability performance, considering the fine-pitch and small bump it contains and the harsh environment it was subjected to. A Weibull plot was drawn based on the Quadrant 2 AATC reliability data, as shown in Figure 4-13. The failures are wear-out failures, as indicated by a β value of 3.5051. The Weibull life of the AATC tested Quadrant 2 samples is 2442 cycles. The sample correlation coefficient is 0.9674, which suggests a high level of confidence that the charted Weibull distribution fits the reliability data well.

Figure 4-14 shows the C-SAM image at 0 and 5000 cycles of Quadrant 3 first-level assemblies, with underfill 3 applied by the needle dispensing method. Figure 4-15 shows the C-SAM image at 0 and 4000 cycles taken on test samples with underfill 4 dispensed by the jetting method. According to the C-SAM images, no underfill delamination was observed in samples underfilled by either underfill dispensing method. However, it was observed in the C-SAM images that samples underfilled with the jet dispensing method show fewer voids and better volume control over the underfill bleeding out than samples underfilled with the needle dispensing method. The reason for this is that the jetting method is capable of dispensing very tiny droplets so that no air would be trapped inside the droplet. For the underfill used in this study, a 0.025mg/dot dispensing volume was achieved by the jetting method.

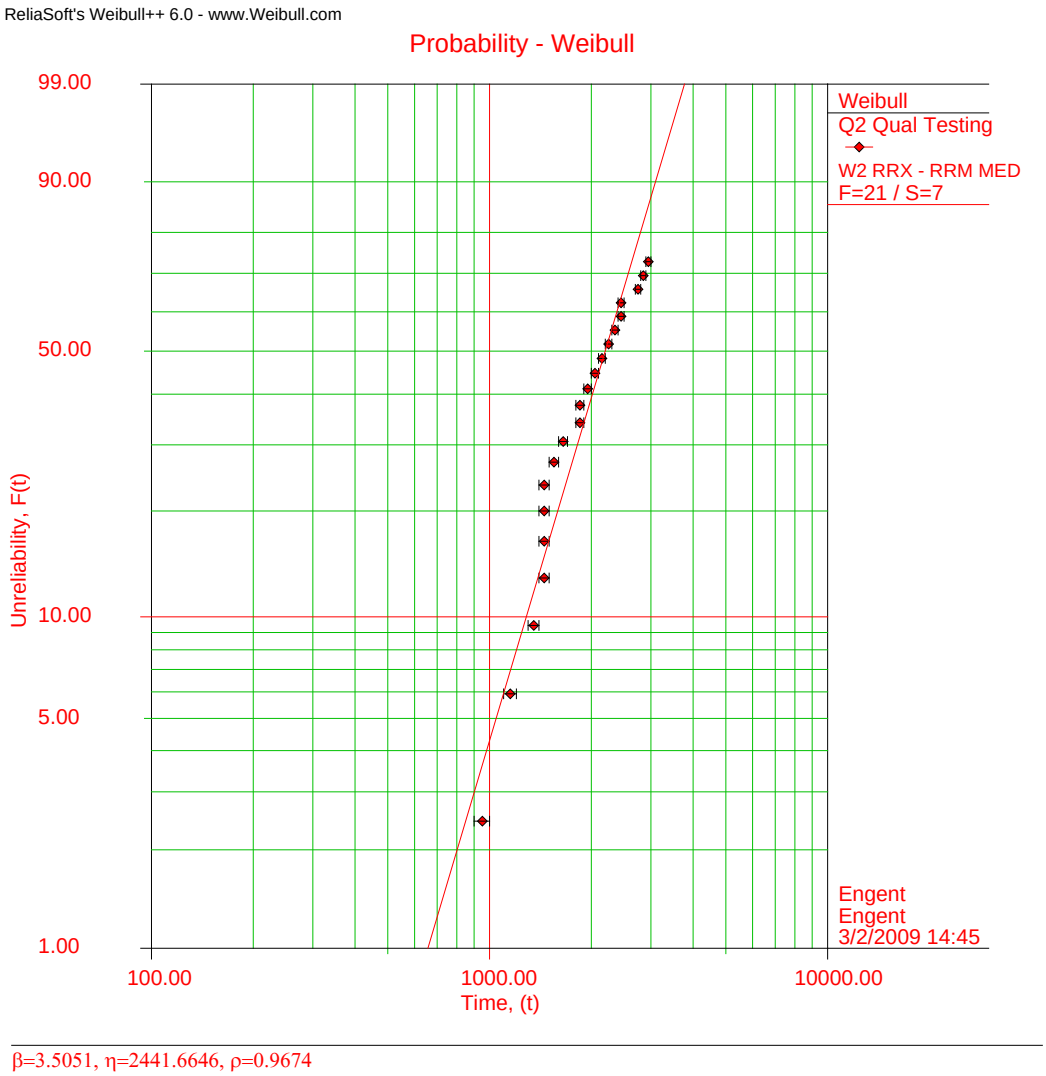


Figure 4-13 Weibull Distribution of Q2 AATC Testing Data

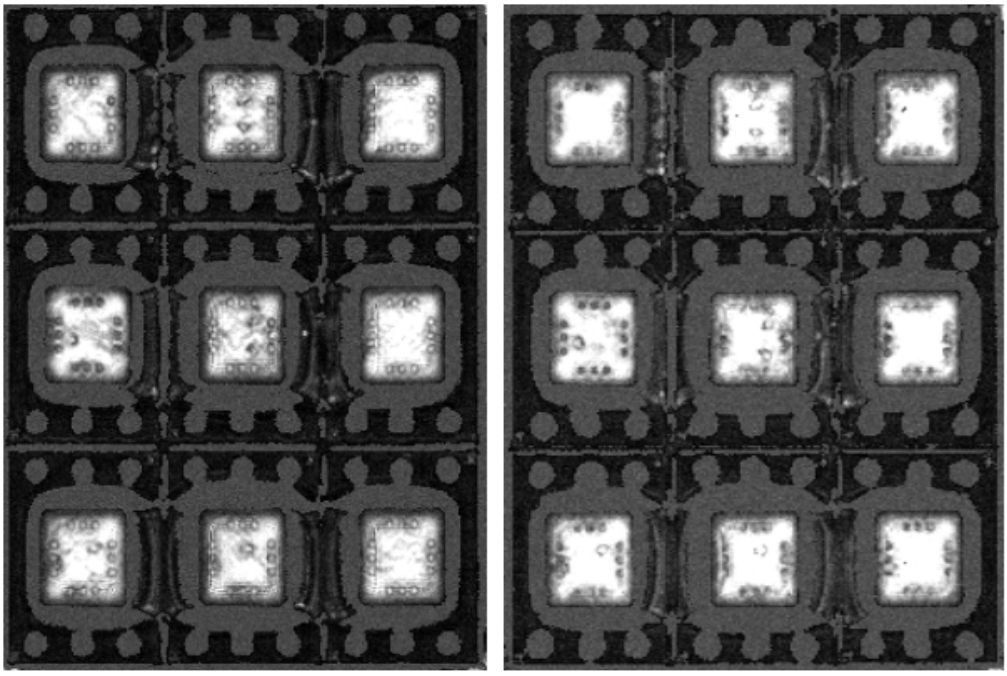


Figure 4-14 C-SAM of Quadrant 3 at 0 cycle (left) and after 5000 cycles (right), Underfill 3 Needle Dispensed

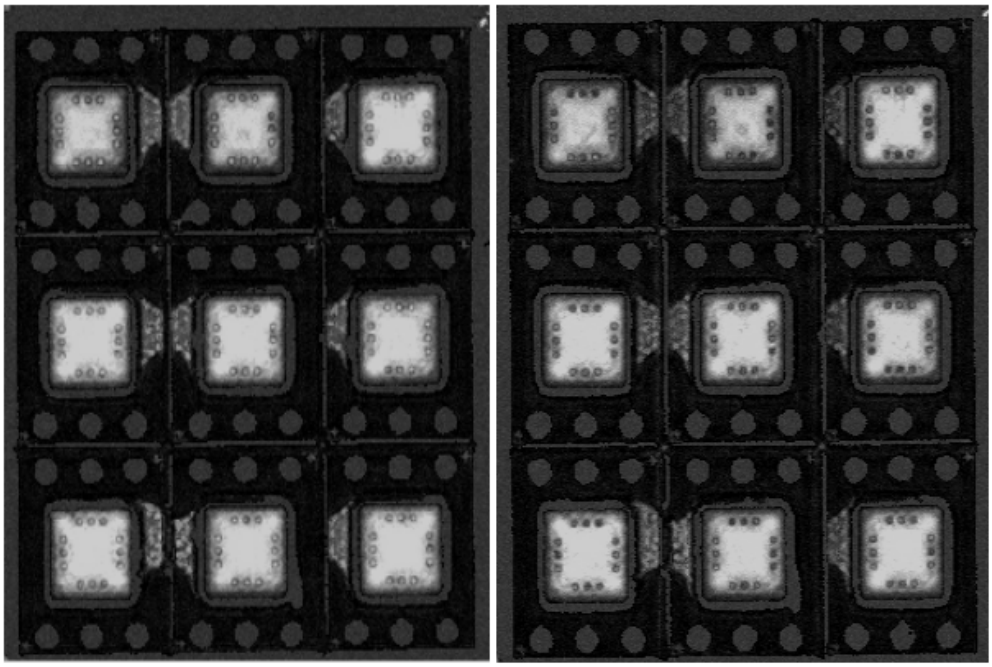


Figure 4-15 C-SAM of Quadrant 3 at 0 cycle (left) and after 4000 cycles (right), Underfill 4 Jet Dispensed

Cross section was performed on a sampling of Quadrant 2 assemblies that failed AATC testing. Figure 4-16 shows an SEM image of a failed Quadrant 2 solder bump. This result and others confirm that solder joint fatigue during cycling was the primary failure mode for Quadrant 2 assemblies that failed in AATC testing.

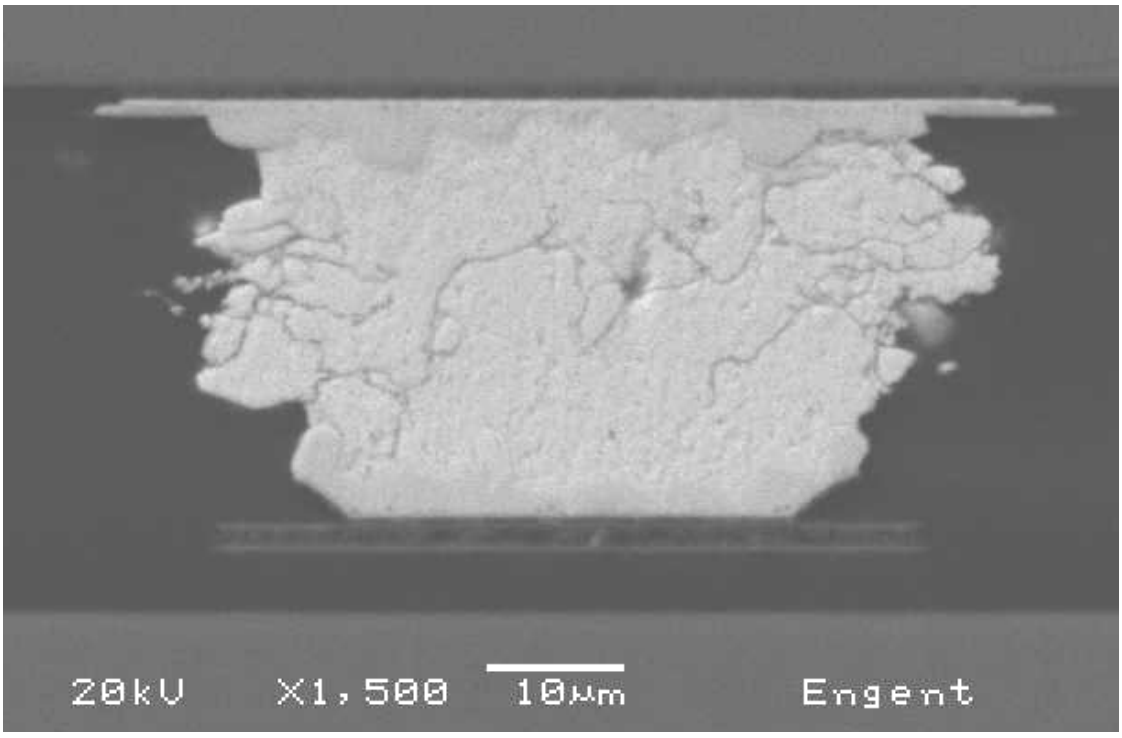


Figure 4-16 Cross Section of Failed Quadrant 2 Bump

For the Quadrant 3 samples which were underfilled with the needle dispensing method, there was only one failure out of 89 samples after 5000 cycles of AATC testing, as the resistance across the daisy chain rose from 1.32 ohm to 12 ohm. Figure 4-17 shows the X-Ray image taken on the failed sample (left) and a good sample (right) after 5000 cycles. Two solder joints from the image on the left side were found to have deformation, as pointed out by the blue arrows. A cross section of the failed sample was performed and

attention was focused on the deformed solder joints shown by the X-Ray. The cross section of one of the two deformed solder joints is shown in Figure 4-18.

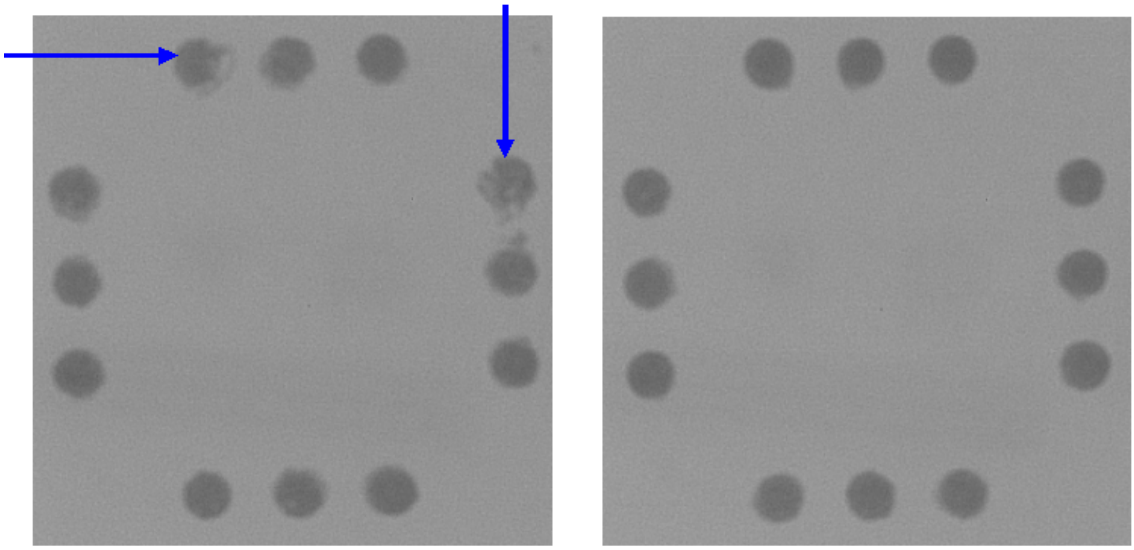


Figure 4-17 X-Ray of the Failed Quadrant 3 Sample (Left) and a Good Quadrant 3 Sample (Right) After 5000 Cycles

It was found in Figure 4-18 that the cracks were not severe enough to cause an open solder joint but the resistance across the daisy chain increased due to the cracks. Note that in Figure 4-16 very little underfill material was observed at the surrounding area of the failed solder joint. The absence of underfill accounts for the excessive stress applied to the solder joint without proper support from the underfill material. The empty space surrounding the solder joint also allows the solder to deform and extrude as shown in Figure 4-18.

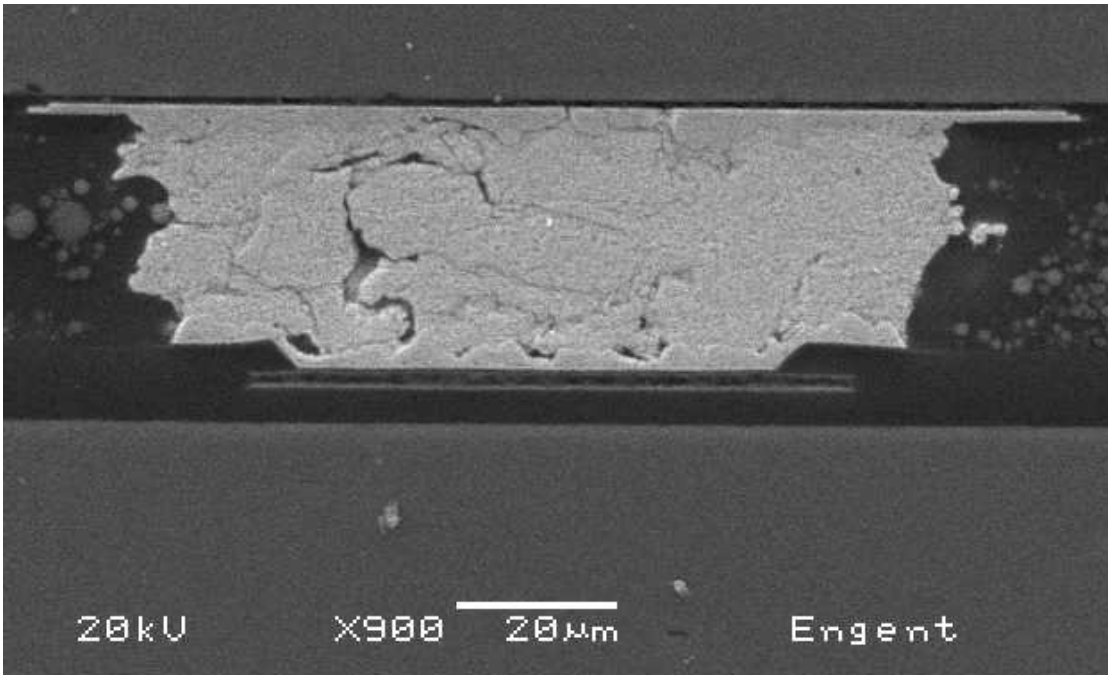


Figure 4-18 Cross Section of Failed Quadrant 3 Solder Joint After 5000 Cycles

Figure 4-19 shows the cross section view of a good solder joint that survived 5000 cycles of AATC testing. Note that more underfill filler particles were identified in the area surrounding the bump in Figure 4-19 than in Figure 4-18. This can explain the reason that the solder joint in Figure 4-19 shows better reliability performance than that in Figure 4-18.

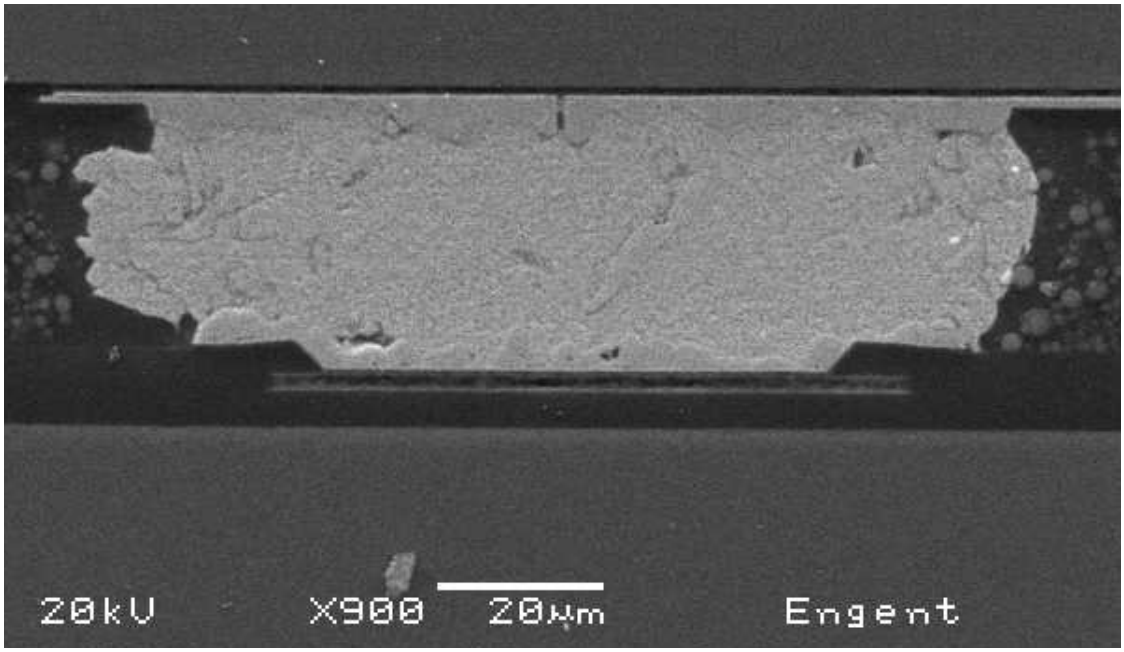


Figure 4-19 Cross Section of a Good Quadrant 3 Solder Joint After 5000 Cycles

4.4 Summary

This chapter explored the application of capillary underfill for the 3D-WLCSP first-level assembly. Although the first-level assembly is a silicon-on-silicon attachment with the flip chip process, underfill is applied to overcome the mechanical stress induced by wafer dicing, second-level assembly pick and place activities, as well as by product handling. Two package design layouts, three flux materials and four underfill materials were evaluated. Both positive displacement pump dispensing and jet dispensing were demonstrated for 3D-WLCSP first-level assembly processing. Tighter control of the underfill volume and a lesser degree of bleed-out onto the WLCSP second-level solder balls was achieved with the jetted underfill technology.

Reliability results indicated that the underfill selection and the package layout had significant impact on the reliability of the 3D-WLCSP first-level package. Flux selection

is not a significant factor toward reliability. Three of the capillary underfill and flux combinations (of the 12 evaluated) showed robust LLTS reliability and none of the 12 material combinations exhibited failure during unbiased autoclave testing.

Minimal delamination was observed in the flip chip assemblies throughout the LLTS and unbiased autoclave reliability tests for the three high performing material sets.

Based on the flux/underfill compatibility study, a more comprehensive package qualification testing at first-level was conducted to evaluate the long-term reliability performance of 3D-WLCSP first-level assembly. Quadrant 3 assemblies with underfill 3 needle dispensed show only one failure up to 5000 cycles of AATC testing. No failure was observed on Quadrant 3 samples with underfill 4 jet dispensed during 4000 AATC testing. Quadrant 2 assemblies with 85-micron pitch show satisfactory reliability performance with the first failure observed after 900 cycles. Solder joint fatigue induced by thermal cycling is the root cause of Quadrant 2 and Quadrant 3 failures. Minimal underfill delamination was observed in the assemblies throughout the AATC testing.

This part of work has demonstrated that the first level package of the novel 3D-WLCSP technology can undertake harsh thermal environment with long-term, high reliability performance and thus to be a qualified package architecture for the Pb-free 3D die-to-wafer integration solution.

CHAPTER 5 3D-WLCSP FIRST-LEVEL NO-FLOW UNDERFILL FLIP CHIP ASSEMBLY PROCESS DEVELOPMENT

5.1 Introduction

Underfill encroachment on the solder balls of Quadrant 1 PA wafer was observed after first-level assembly with capillary underfill assembly process. Figure 5-1 shows the comparison of the Quadrant 1 and Quadrant 4 first-level packages after capillary underfill. The underfill was found to bleed-out and overflow onto the PA solder balls in Quadrant 1 packages. The reason for this phenomenon is that the Quadrant 1 PA has a perimeter bumped structure with the flip chip die in close proximity. This structure creates difficulty for the first-level underfill application on Quadrant 1 PA wafer as the underfill dispensed, either by positive displacement pump or jet dispenser, tends to overflow onto the CSP balls close-by. This creates a process difficulty for the second-level assembly, when the first-level packaged WLCSP is surface mounted to the PCB with interconnection provided by the solder balls. The underfill covering solder balls acts as a barrier that prevents solder balls from wetting on the PCB pads. This greatly affects the assembly yield and in the long run, the reliability of the whole 3D-WLCSP second-level package.

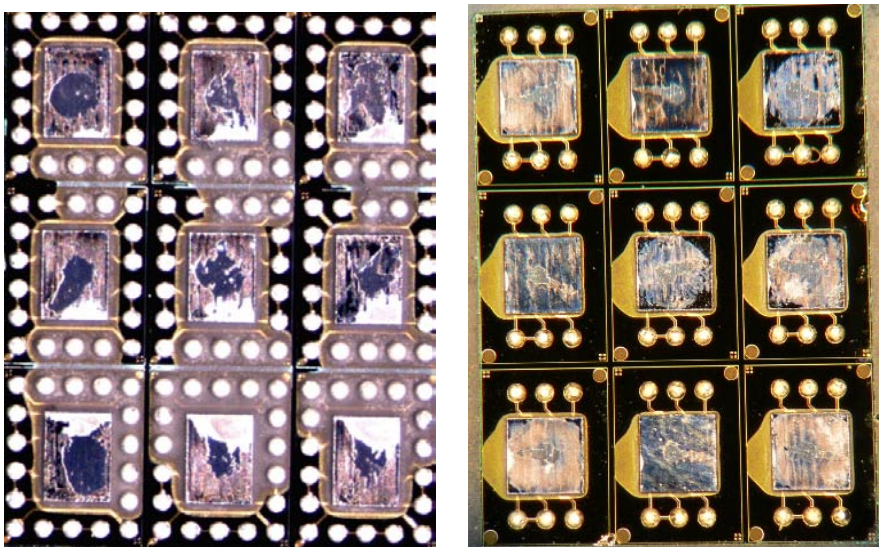


Figure 5-1 3D-WLCSP First-level Assembly with Capillary Underfill, Quadrant 1 (left), Quadrant 4 (right)

As discussed in Chapter 1.2.4, no-flow underfill can dramatically reduce the process time and the cost per package, due to the reduction in the number of process steps as well as elimination of the dispenser and cure oven that would otherwise be necessary for the standard capillary underfill process. Furthermore, as the underfill is applied before the chip placement, the dispensed volume and location can be carefully controlled so that no excessive underfill material will encroach on the PA WLCSP solder balls.

Figure 5-2 describes the 3D-WLCSP first-level assembly process with no-flow underfill. The no-flow underfill is applied onto each individual WLCSP substrate on the wafer before chip placement. The underfill cure and solder joint formation occur together during the wafer reflow. Compared with the capillary underfill process, no-flow underfill best meets the requirements of the flip chip to wafer integration as the process steps are simplified and no dedicated underfill curing oven is needed. By this method, the overall

throughput can be greatly improved and the cost per package can be substantially reduced.

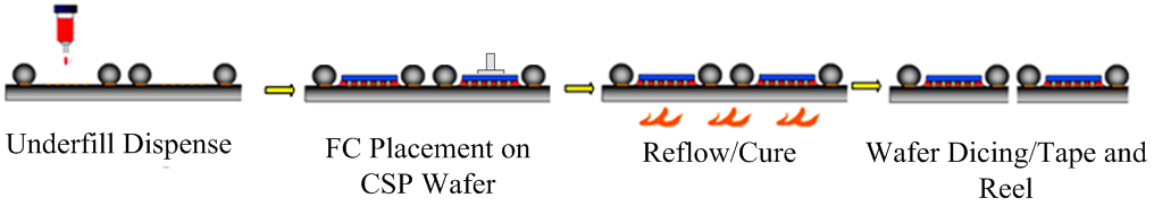


Figure 5-2 3D-WLCSP First-level Assembly No-flow Underfill Process

This chapter introduces the development of the no-flow underfill assembly process for the 3D-WLCSP first-level package. The goal of this research was to develop a robust no-flow underfill process for the fine-pitch flip chip silicon-to-silicon wafer level integration. Challenges addressed include the no-flow underfill reflow profile study, underfill volume study, chip floating control, underfill voiding reduction, and yield improvement. Also, different no-flow underfill candidates were investigated for the selection of best performing no-flow underfill material.

5.2 Experimental Approach and Results

The Quadrant 1 test vehicle was used in this study, as the Quadrant 1 test vehicle had the worst underfill encroachment issue (underfill bleeding out onto CSP solder balls) among the four test vehicle layouts. A 2x2 array PA panel containing 4 WLCSP substrates was utilized in the study to mimic the wafer level packaging structure and facilitate the process development, as shown in Figure 5-3.

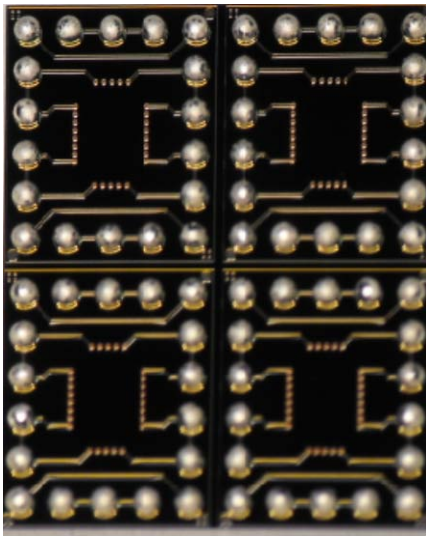


Figure 5-3 Quadrant 1 2x2 Array Test Vehicle Substrate Tile

Two no-flow underfill materials (NF1 and NF2) were selected for this study. The NF2 has lower Tg and Viscosity than NF1. The CTE of both materials is around 80ppm/°C. The selection of NF1 and NF2 was based on manufacturer recommendation and previous project application experience.

5.2.1 Soak Zone Study

A soak zone is needed to activate the fluxing agent in the no-flow underfill material so that the oxidation on the solder bumps and pads can be effectively removed. One critical issue for no-flow underfill application is that the underfill should not cure before the solder begins to wet on the pads. Otherwise, the hardened underfill material will prevent the solder from wetting and collapsing. As a result, no solder joint interconnection can be formed. Therefore, it is important to understand the soak zone of

the no-flow underfill material so that the oxidation can be effectively removed while the underfill remains uncured.

To achieve this goal, a drop of no-flow material was put on a glass slide, which was placed on top of a hot plate. The no-flow underfill was monitored under the microscope with different temperatures applied at the hot plate. A sharp needle was used to probe the underfill in order to check if the material maintained a liquid status during the dwell time. The hot plate study on NF 1 is shown in Figure 5-4.

Figure 5-4 No-flow 1 Soak Zone Study on Glass Slide (Left-25°C, Middle-120°C, Right-160°C for 2 minutes)

No underfill curing activities were identified in the Figure 5-4 left and middle images, as the underfill droplet maintained liquid status in both scenarios. However, NF1 was found cured after 2 minutes dwell time at 160°C, as shown in the right image of Figure 5-4. This was also indicated by the color change of the underfill droplet. Based on this study, it was found that NF1 starts to cure at the temperature as low as 150°C. In comparison, NF2 does not start to cure before 190°C. Therefore, a common reflow profile containing a soak zone of less than 140°C was selected initially for both no-flow underfill materials.

5.2.2 Solder Wettability Study

This study was to help understand the reflow profile condition in which the Pb-free solder can wet on the copper pads with no-flow underfill applied. SAC305 balls and copper clad were used to facilitate the study. The reason for using the copper clad is to save the test vehicles. A 28mil diameter SAC305 solder ball was placed on a copper clad, with one drop of no-flow underfill applied, as shown in Figure 5-5. The copper clad was then reflowed on a carrier in the convection reflow oven to see if the solder could wet on the copper under certain reflow profiles.

Figure 5-5 Solder Wettability Test with SAC305 Solder Ball on Copper Clad

Figure 5-6 and Figure 5-7 show the solder wetting condition after the copper clad was reflowed under different reflow profiles. The left image in Figure 6-6 shows that the solder did not wet on the pad. The non-wet condition was confirmed by the shear test, as shown in the right image of Figure 5-6. No proper intermetallic compound was formed and the solder ball did not collapse. In comparison, Figure 5-7 shows the solder ball

collapsed and wet on the copper clad. The wetting condition was confirmed by the shear test which showed the area of the intermetallic compound.

Based on the solder wettability testing, it was found that the reflow profile process window of NF2 is much wider than NF1, as solder wets much easier with NF2 than with NF1 under the same profile.

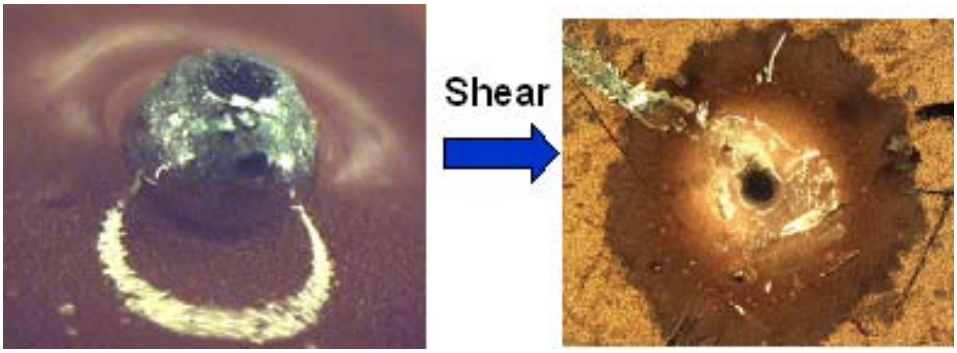


Figure 5-6 Solder Wettability Testing (Non-wet)

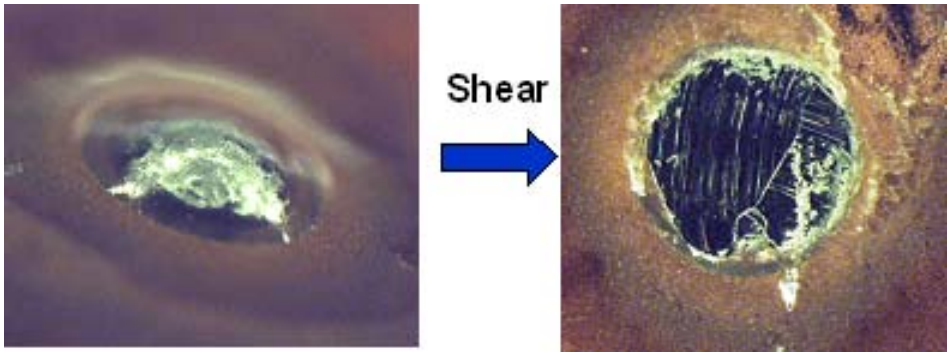


Figure 5-7 Solder Wettability Testing (Wet)

Based on the soak zone study and solder wettability study, a reflow profile of 35 seconds soak at 120°C -140°C, peak temperature of 235°C and reflow time of 40 seconds was found to wet the Pb-free solder with both NF1 and NF2. This profile is shown in Figure 5-8. A fast ramp after soak is desired to make sure the solder starts to melt and wet on the pad before the underfill cures.

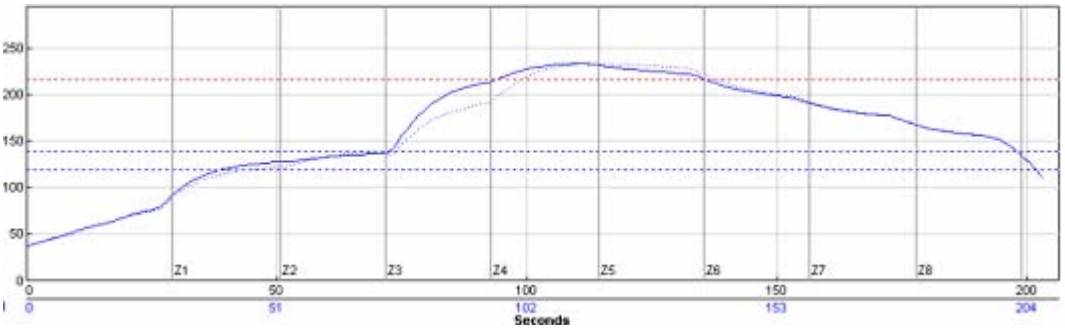


Figure 5-8 Initial Reflow Profile for NF1 and NF2

5.2.3 No-flow Underfill Dispense Volume Study

The no-flow underfill chip floating phenomena was discussed by Thorpe and Baldwin [50]. It was reported that chips may float if too much underfill is applied onto the substrate. In this case, no interconnection can be achieved. This is especially true in the case of fine-pitch flip chip assembly due to the 100μm thin profile of the flip chip. On the other hand, a too small amount of no-flow underfill dispensed may cause the lack of underfill fillet which will compromise the package long term reliability performance. Therefore it is vital to understand and control the proper underfill amount dispensed.

An Asymtek DJ-9000 jetting dispenser was utilized in this study due to its precise control over the underfill volume and its capability of jetting very small underfill droplets (0.05mg/dot in this study). After underfill dispense, the flip chips were picked and placed by the ESEC Micron 2 chip placement machine. Figure 5-9, Figure 5-10 and Figure 5-11 show the packages assembled with underfill amounts at 0.05mg, 0.15mg and 0.3mg, respectively. The measurement of the resistance across daisy chains on samples shown in Figure 5-9 and Figure 5-10 indicated interconnection achieved after reflow in both scenarios.

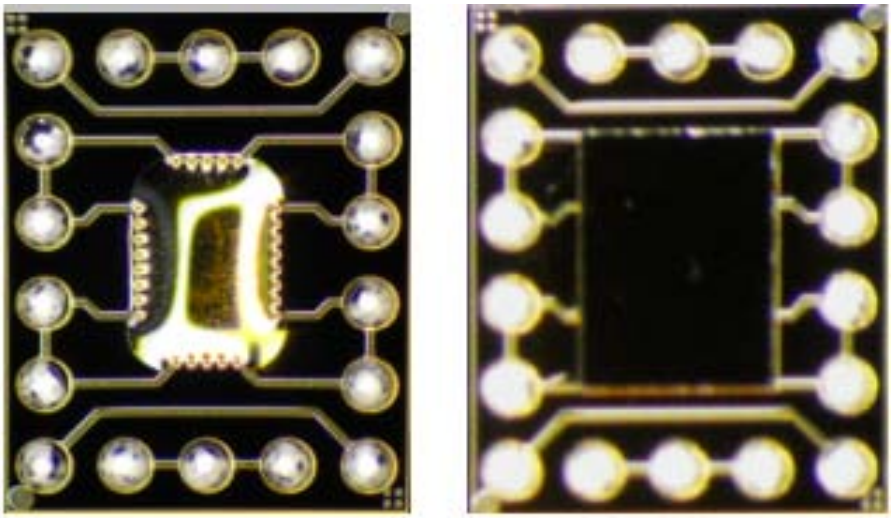


Figure 5-9 0.05mg Underfill after Dispense (Left) and after Reflow (Right)

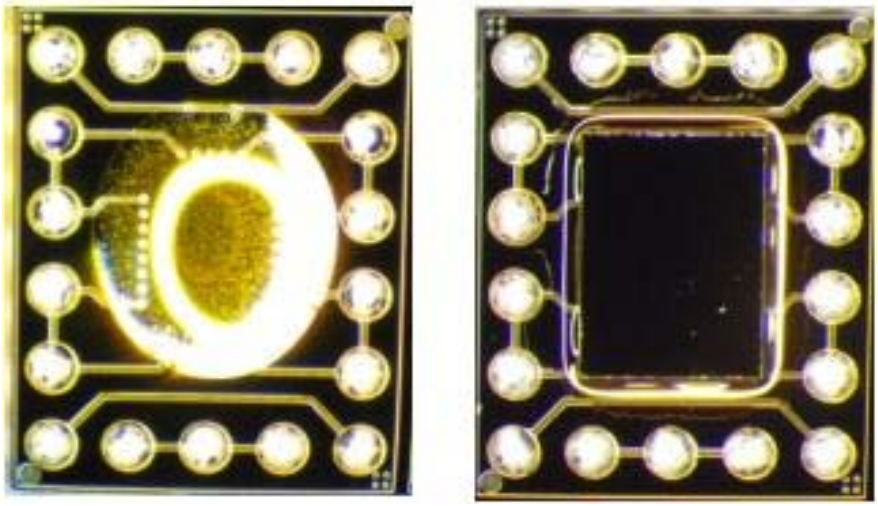


Figure 5-10 0.15mg Underfill after Dispense (Left) and after Reflow (Right)

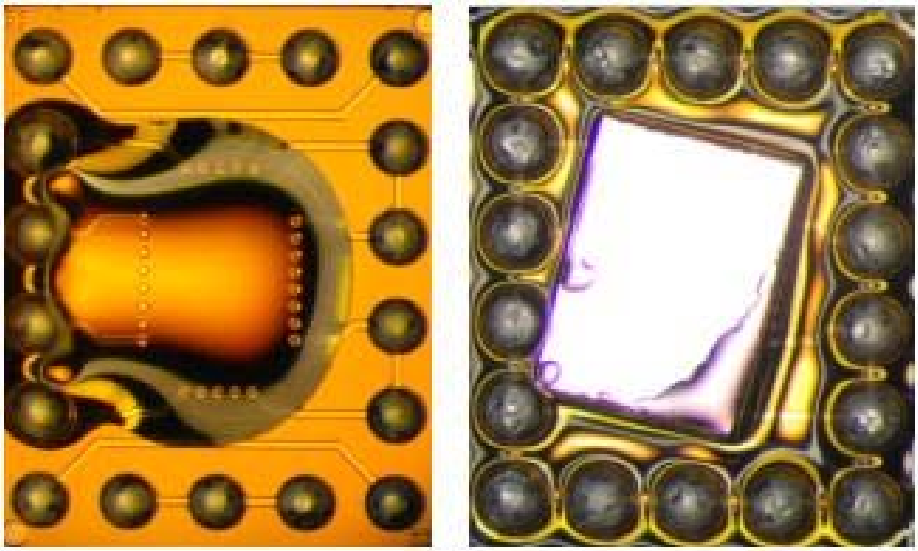


Figure 5-11 0.3mg Underfill after Dispense (Left) and after Reflow (Right)

No underfill fillet was observed after reflow in Figure 5-9, indicating that the underfill amount is not enough. Figure 5-11 shows the case in which excessive underfill amount can cause the chip to float so that no interconnection can be achieved. Figure 5-10 shows the proper underfill fillet formed after reflow. This indicates that the underfill amount dispensed is correct. Based on the underfill volume study, it was found that 0.15mg was the proper volume to be dispensed for the Quadrant 1 first-level no-flow underfill assembly.

5.2.4 Initial Assembly Evaluation

Initial assembly was made on samples assembled with both NF1 and NF2, following the reflow profile shown in Figure 6-8. The underfill was dispensed by the Asymtek DJ-9000 jetting dispenser with a fixed dispense mass of 0.15mg. A total of 20 NF1 samples and 20 NF2 samples were assembled. For the initial assembly, Quadrant 1

first-level packages assembled with NF1 achieved 75% yield while packages assembled with NF2 had 100% yield.

Cross sections were performed on packages assembled with NF1 and NF2, as shown in Figure 5-12 and Figure 5-13. It was found that the solder joint formed with NF2 had better quality than the solder joint formed with NF1, as the NF2 case shows much better intermetallic compound (IMC) formation. The lack of proper IMC accounts for the yield loss in the NF1 assembled samples.

C-mode scanning acoustic microcopy (C-SAM) was utilized to check the underfill voiding condition, as shown in Figure 5-14. The NF1 samples show less underfill voiding than NF2 samples.

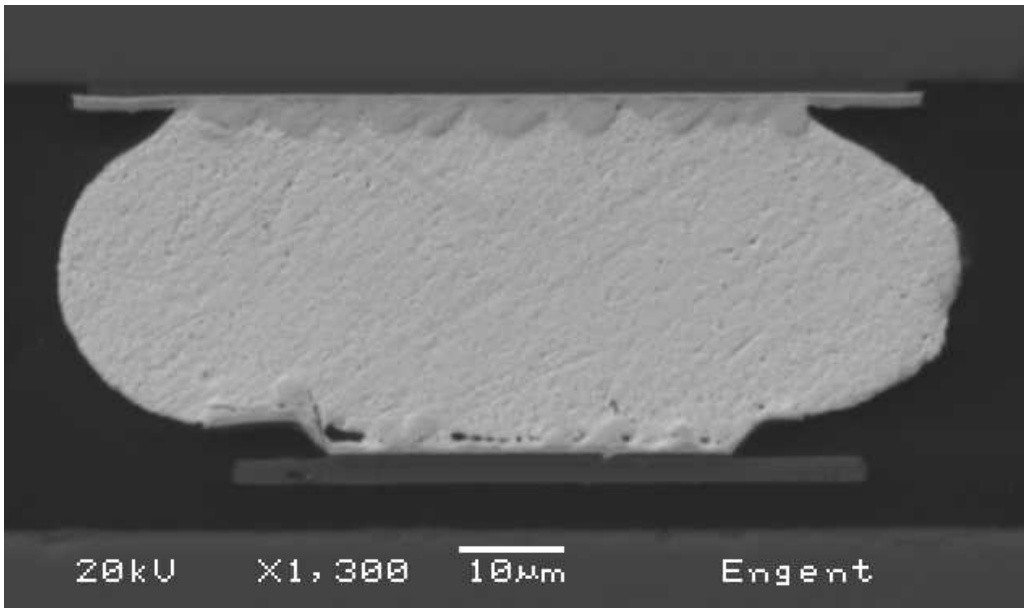


Figure 5-12 Cross Section of a Solder Joint Formed by NF1

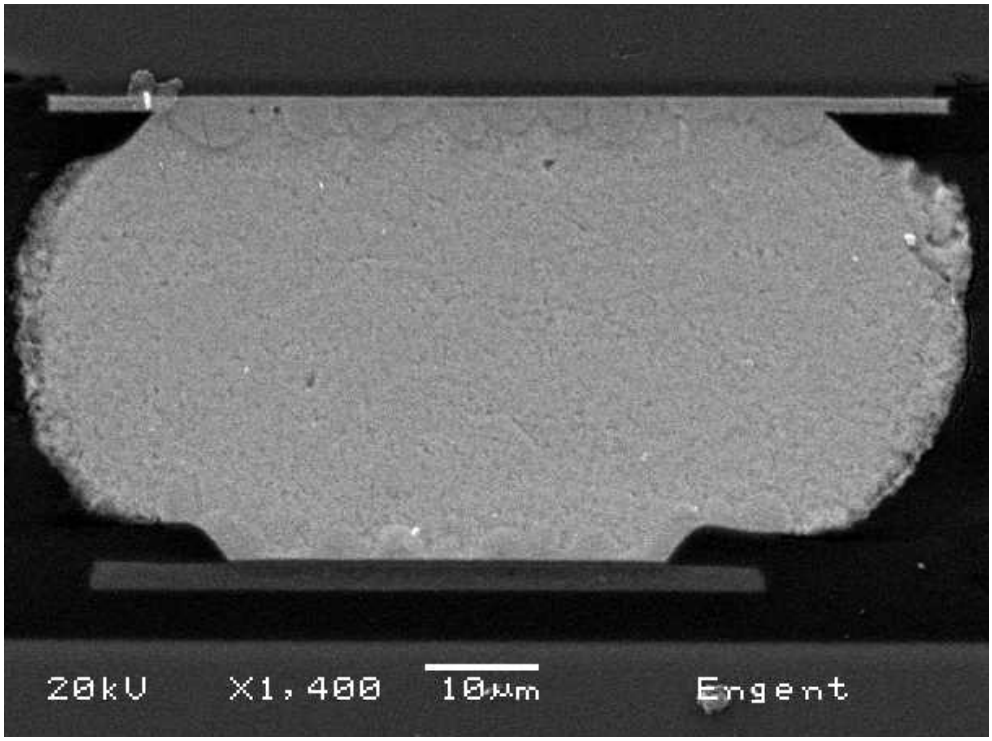


Figure 5-13 Cross Section of a Solder Joint Formed by NF2

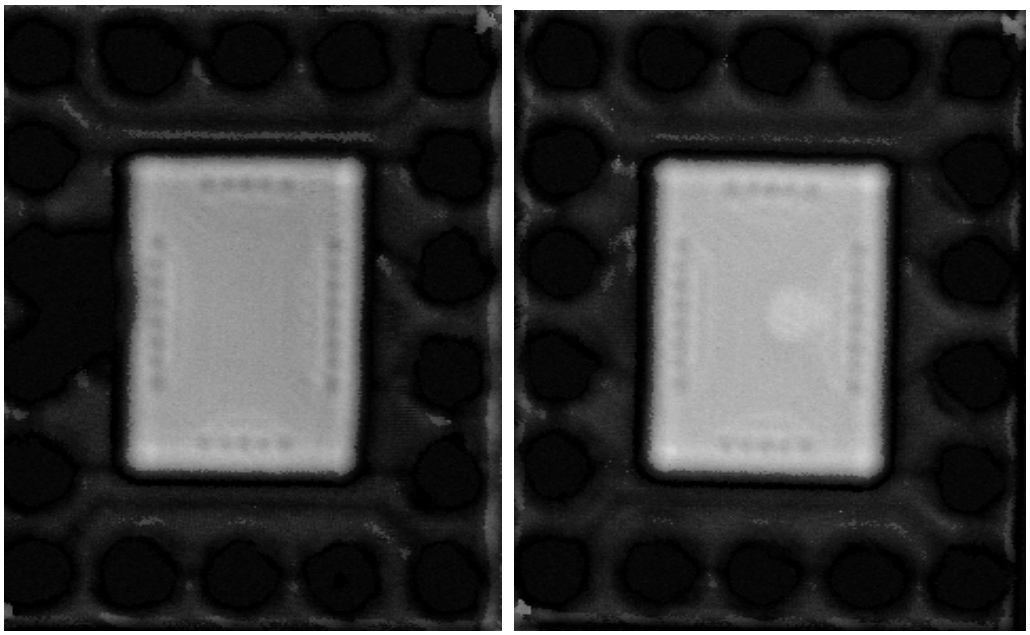


Figure 5-14 C-SAM Images of Package Assembled with NF1 (Left) and NF2 (Right)

5.2.5 No-flow Underfill Reflow Process Improvement

The initial assembly revealed that samples built with NF1 had yield loss due to inadequate intermetallic compound formation. One possible reason is that the peak temperature in the initial reflow profile was too high for NF1 so that the underfill cured before the formation of intermetallic compound. The other possible reason is that the reflow time was not long enough for the formation of a proper intermetallic compound. Therefore, the reflow profile was modified accordingly, with lower peak temperature and longer time above liquidus. The comparison of the initial profile and the improved profile are shown in Figure 5-15.

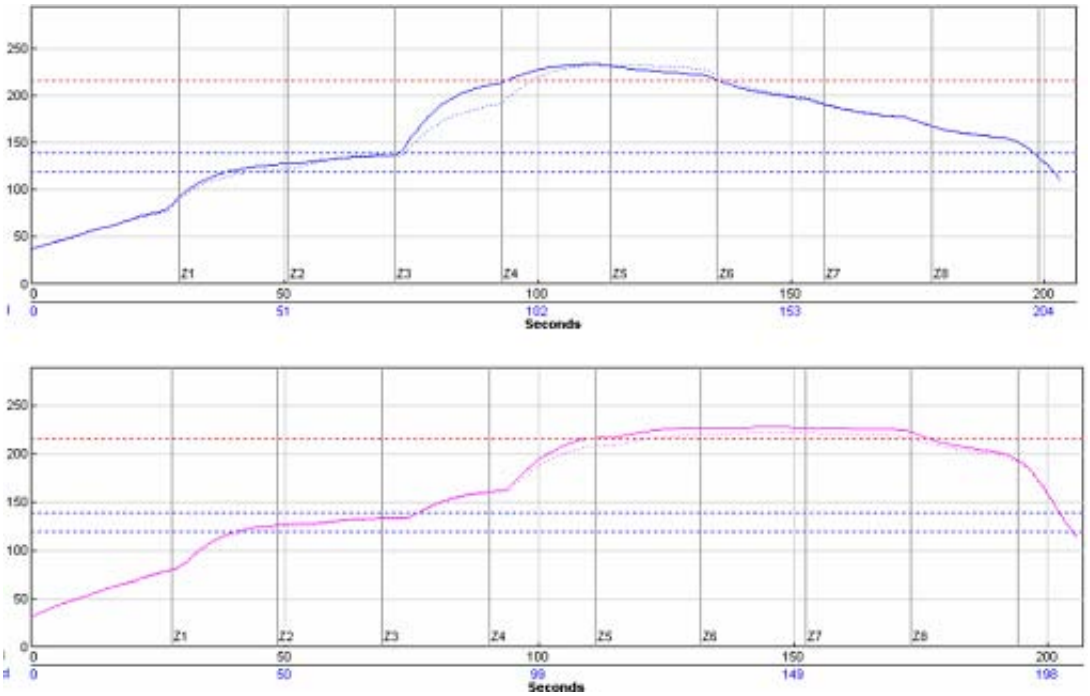


Figure 5-15 Comparison of Initial Reflow Profile (upper) and Improved Reflow Profile (lower) for NF1

Table 5-1 shows the parameter comparison between the initial reflow profile and the improved reflow profile for NF1. As a result, the yield was improved from 75% to 100% on a sample of 20 assembled units.

Table 5-1 NF1 Initial and Improved Reflow Profile Parameter Comparison

NF1	Soak Zone(°C)	Soak Time(s)	Peak Temperature(°C)	Time Above Liquidus(s)	Yield
Initial	120-140	35	235	40	75%
Improved	120-140	35	229	65	100%

The initial NF2 assembled samples show underfill voids in certain packages, as shown in Figure 6-14. According to Lee and Baldwin [22], changing the soak zone and soak time can reduce the no-flow underfill voids. As shown in the previous soak zone study, NF2 does not cure before 190°C. Thus the NF2 reflow profile was modified with a new soak zone of 150°C -180°C and a longer soak time of 110 seconds. The time above liquidus was also increased to make sure the voids could be effectively driven out during reflow. The initial and improved NF2 reflow profiles are shown in Figure 5-16.

Table 5-2 shows the comparison between the initial reflow profile and the improved reflow profile for NF2. A C-SAM image was taken on samples built with NF2 under the improved reflow profile, as shown in Figure 5-17. Much less underfill voiding was identified on samples built with improved NF2 reflow profile.

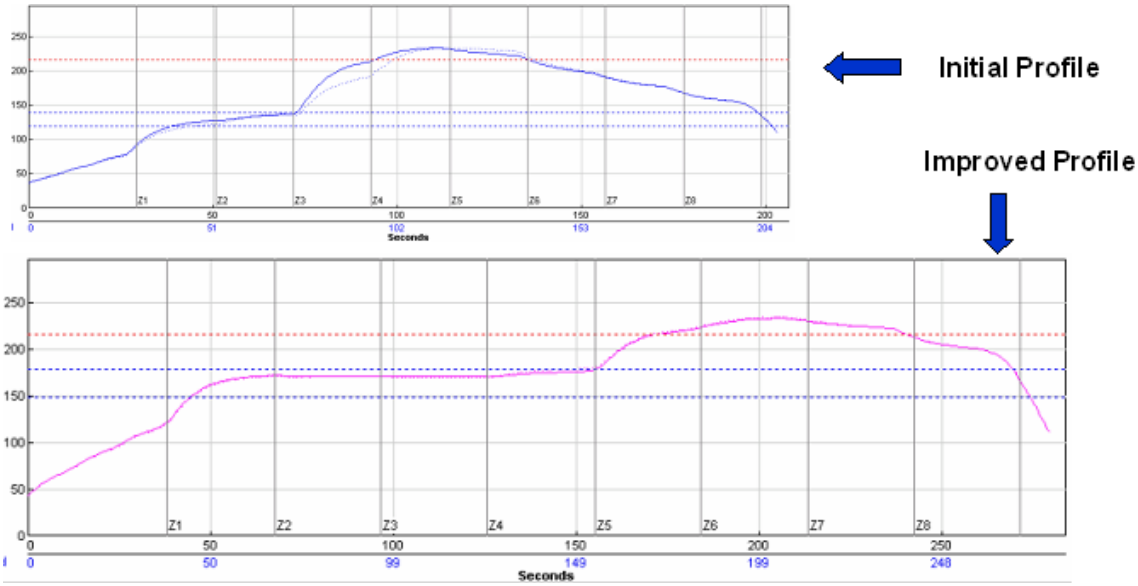


Figure 5-16 Comparison of Initial Reflow Profile (upper) and Improved Reflow Profile (lower) for NF1

Table 5-2 NF2 Initial and Improved Reflow Profile Parameter Comparison

NF2	Soak Zone(°C)	Soak Time(s)	Peak Temperature(°C)	Time Above Liquidus(s)	Yield
Initial	120-140	35	235	40	100%
Improved	150-180	110	235	70	100%

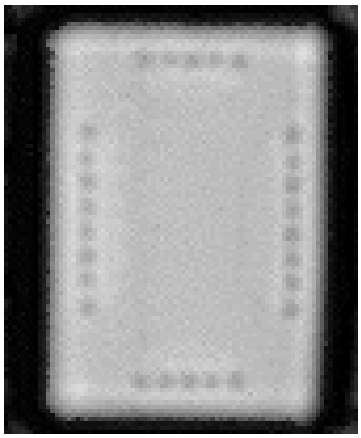


Figure 5-17 C-SAM image on package assembled with the improved reflow profile on NF2

5.3 Summary

This chapter introduced the development of a fine-pitch flip chip to CSP wafer level integration with the application of no-flow underfill. Two no-flow underfill materials were investigated with NF2 showing a wider process window and better performance on improving solder wettability.

Underfill volume dispensed plays an important role in terms of the underfill fillet formation and chip floating phenomena. A too small underfill volume dispensed will cause the lack of proper underfill fillet. This will compromise the package's long-term reliability. Too much underfill applied may cause the chip to float so that no interconnection can be achieved. This will induce the 3D-WLCSP wafer level assembly yield loss. An underfill volume of 0.15 mg was found to be proper for the formation of fillet and prevention of chip floating for the Quadrant 1 test vehicle in this study.

Soak zone and solder wettability studies were performed for the initial reflow profile set up. Initial samples built with NF1 show yield loss due to inadequate formation of an intermetallic compound. An improved reflow profile with lower peak temperature and longer time above liquidus was designed to solve this problem. Initial samples built with NF2 show underfill voids after reflow. The underfill voiding issue was solved by the improved reflow profile that has higher soak zone temperature, longer soak time and longer time above liquidus. Both NF1 and NF2 achieved 100% yield after the reflow profile optimization.

This part of work demonstrated that the no-flow underfill is an applicable method to improve the 3D-WLCSP first-level assembly throughput with high assembly yield.

CHAPTER 6 3D-WLCSP WAFER LEVEL ASSEMBLY SCALE UP STUDY

6.1 Introduction

Previous chapters have demonstrated that the 3D-WLCSP first-level assembly can achieve 100% yield with flip chip assembly process. The previous research and development work on the 3D-WLCSP first-level assembly was conducted with the test vehicles of 2x2 and 3x3 array tiles, as shown in Figure 4-3, 4-4 and 5-3. The reason for using the tiles was to facilitate the development work and save on test vehicle wafers.

With the high yield and reliable flip chip assembly process developed previously, a wafer level scale-up study is carried out in this chapter. The test vehicle WLCSP wafer has a total of 1,876 WLCSP substrates fabricated on it and is divided into four different quadrants with different layouts, as shown previously in Figure 3-5.

There are several challenges associated with the wafer level assembly compared with the assembly on the 2x2 or 3x3 tiles. The first challenge is the creation of the pick-and-placement program on the wafer scale. For example, only 9 locations need to be input in the ESEC Micron 2 placement program with a 3x3 array test vehicle tile. By contrast, the wafer assembly program requires 1,876 placement locations to be known by the machine in order to process all the pick and placement activities.

The second challenge associated with the wafer level assembly is the placement speed-up. The initial assembly process developed on the 2x2 or 3x3 tiles utilized only 20% of the machine's full speed capability. As the focus was put on the material selection

and reflow profile evaluation, the pick and placement speed was set to a fairly slow level to make sure every chip can be placed accurately. However, with the placement of over 1,800 flip chips, the full travel speed of the machine shall be utilized in order to achieve high throughput with low processing time. Thus it is vital to explore if the high machine travel speed will affect the assembly yield.

The third challenge with the wafer level assembly is the dwell time of the flux material. By the dry run, the estimated time for the ESEC Micron 2 to finish 1,876 times of pick, die vision, dip fluxing and placement is approximately four hours with full machine placement speed. This means the chip placed on the first location shall wait for around four hours before it goes to the reflow oven. Thus it is important to learn if the flux material is still functional after 4 hours of dwell time in the air. If the flux dries out during the long dwell time, the yield loss would be inevitable.

This chapter discusses the wafer level assembly processing challenges and solutions. The goal of this study is to develop a robust die-to-wafer flip chip assembly process for the 3D-WLCSP first-level integration.

6.2 Experimental Approach and Results

6.2.1 Wafer Level Assembly Program Creation

The operation of the ESEC Micron 2 machine requires that the placement location of each individual chips be known to the machine. The most straightforward way is to calculate each individual placement location on the wafer with reference of the wafer fiducials and to key-in the X-Y coordinates of each location into the program. Two disadvantages are associated with this method. The first drawback is the time consumed

for verifying and keying in the X-Y coordinates over 1,800 locations. If the location data were input wrongly, the placement accuracy would be affected with resulting yield loss. The second drawback found with the direct key-in method is that during the X-Y location data entry, the machine gets slower in response to the data keyed in as insufficient memory could be allocated. This results in a very inefficient placement program. Both drawbacks indicate that the direct key-in method is not applicable with the ESEC Micron 2 chip placement machine.

A new way to program the placement activities on the whole wafer was developed. As four different test vehicles were fabricated on the same WLCSP wafer with each test vehicle taking up one quadrant of the wafer, the wafer programming was split into four different programming activities, with each wafer quadrant programmed separately. For example, the first quadrant of the wafer has a total of 506 Q1 WLCSP substrates. The second quadrant of the wafer has a total of 470 Q2 WLCSP substrates. Both the third and fourth wafer quadrants have 450 WLCSP substrates fabricated. Thus the program was created based on the quadrant rather than the whole wafer, as each quadrant wafer has different die and substrate size as well as the layout. The image of quadrant 1 wafer is shown in Figure 6-1.

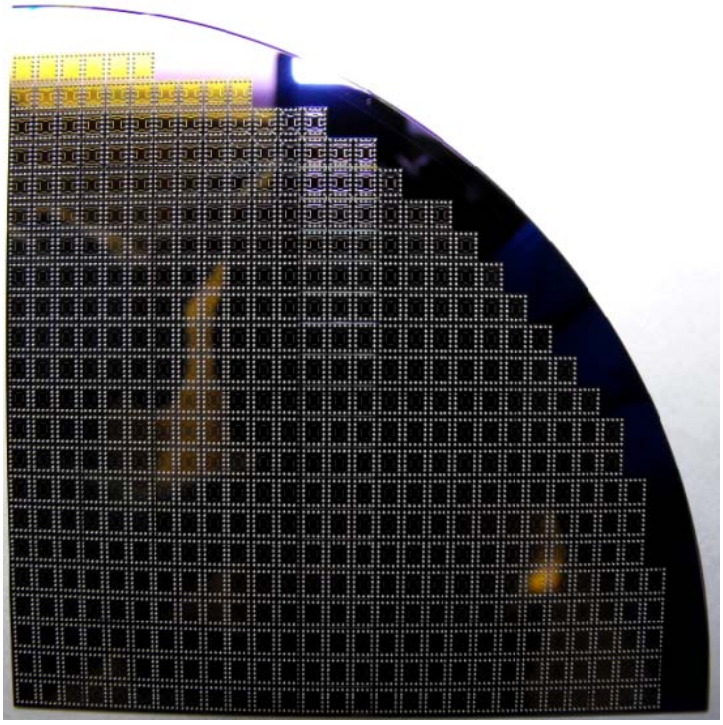


Figure 6-1 WLCSP Wafer Quadrant 1 Test Vehicle

In order to program Quadrant 1 efficiently and save machine response time during program execution, the program for Quadrant 1 was divided into 2 sub-programs. Group "A", as shown in Figure 6-2, was laid out in a rectangular shape, which has a 17x18 array (306 pieces) of CSP substrates. The 17x18 locations in section "A" were programmed by a "step and repeat" programming function provided by the ESEC Micron 2 machine. The first column is programmed (only 18 locations in the first column keyed into the program) and the remaining 17 columns in area "A" were "imaged" from the first column with only the pitch (distance between the first column to the imaged column) input to the program. Therefore, instead of inputting 306 (=18x17) locations to the program, only 35 (=17+18) locations were input to Sub-program 1. This method greatly increased the

program efficiency. For areas B and C on Quadrant 1, the location of each individual substrate was directly input to Sub-program 2, as neither area is in a rectangular shape.

The same programming method was used for Quadrant 2, 3 and 4 partial wafers. In total, there were 8 sub-programs for the whole wafer-scale programming. By dividing the whole wafer into sections and programming according to the layout of each section, the wafer assembly program was accomplished with high efficiency and reduced machine response time.

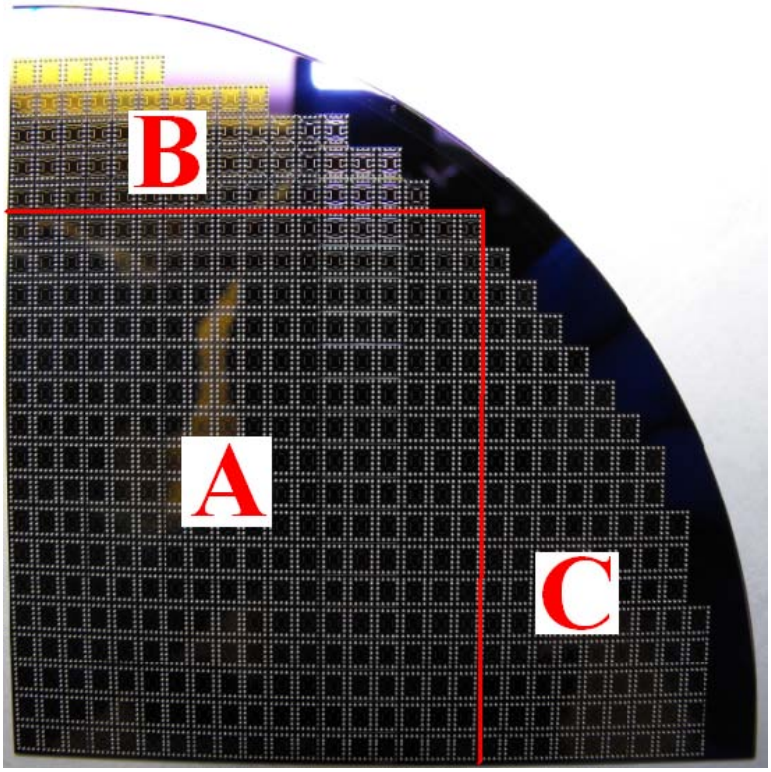


Figure 6-2 Quadrant 1 Area Allocation for the Placement Program

6.2.2 Flux Dwell Time Study

As discussed in Chapter 6.1, the functionality of the flux after dwell time is critical for the wafer level assembly yield. If the flux material becomes dried out and loses its fluxing capability before the reflow of assembled wafer, wafer level assembly yield loss could occur. Therefore, an experiment was conducted to learn how long the flux dwell time could be in order to maintain 100% yield on the 3D-WLCSP first-level assembly.

Quadrant 3 test vehicles were utilized in this study. A total of 45 pieces of Quadrant 3 first-level packages were assembled. Flux B, which had demonstrated high yield and good flux/underfill compatibility, as discussed in Chapter 3, was used. The flux dip height was 45 microns. After the end of every hour of dwell time, 9 pieces of first-level packages were sent to the reflow oven with nitrogen environment, following the reflow profile condition shown in Table 6-1. The reflow parameter settings are based on previous study. The outline of the experiment and results are shown in Table 6-2.

Table 6-1 Flux Dwell Time Reflow Profile Parameter Setting

Reflow Parameter	Ramp Rate	Soak Temp	Soak Time	Time Above Liquidus	Peak Temp
	1.2°C/s	150 ~ 180 °C	50 sec	65 sec	240°C

Table 6-2 Flux B Dwell Time Study

Dwell Time (Hours)	Assembly Yield
1	9/9
2	9/9
3	9/9
4	9/9
5	9/9

The results, as shown in Table 6-2, indicate that the tacky flux material selected for the 3D-WLCSP wafer level assembly has a long dwell time that will suffice the dwell time needed for the wafer level assembly process.

6.2.3 Initial Wafer Level Assembly Evaluation

The initial wafer level assembly was conducted after the creation of the program and the study of the flux dwell time. The reflow profile parameters are shown in Table 6-1. As discussed previously, the program was created based on each wafer quadrant. Thus the initial assembly was also conducted on the quadrant wafers. The full machine speed was utilized to evaluate the machine capability for a high-speed, high throughput wafer level assembly requirement. The images of each quadrant wafer after assembly are shown in Figure 6-3.

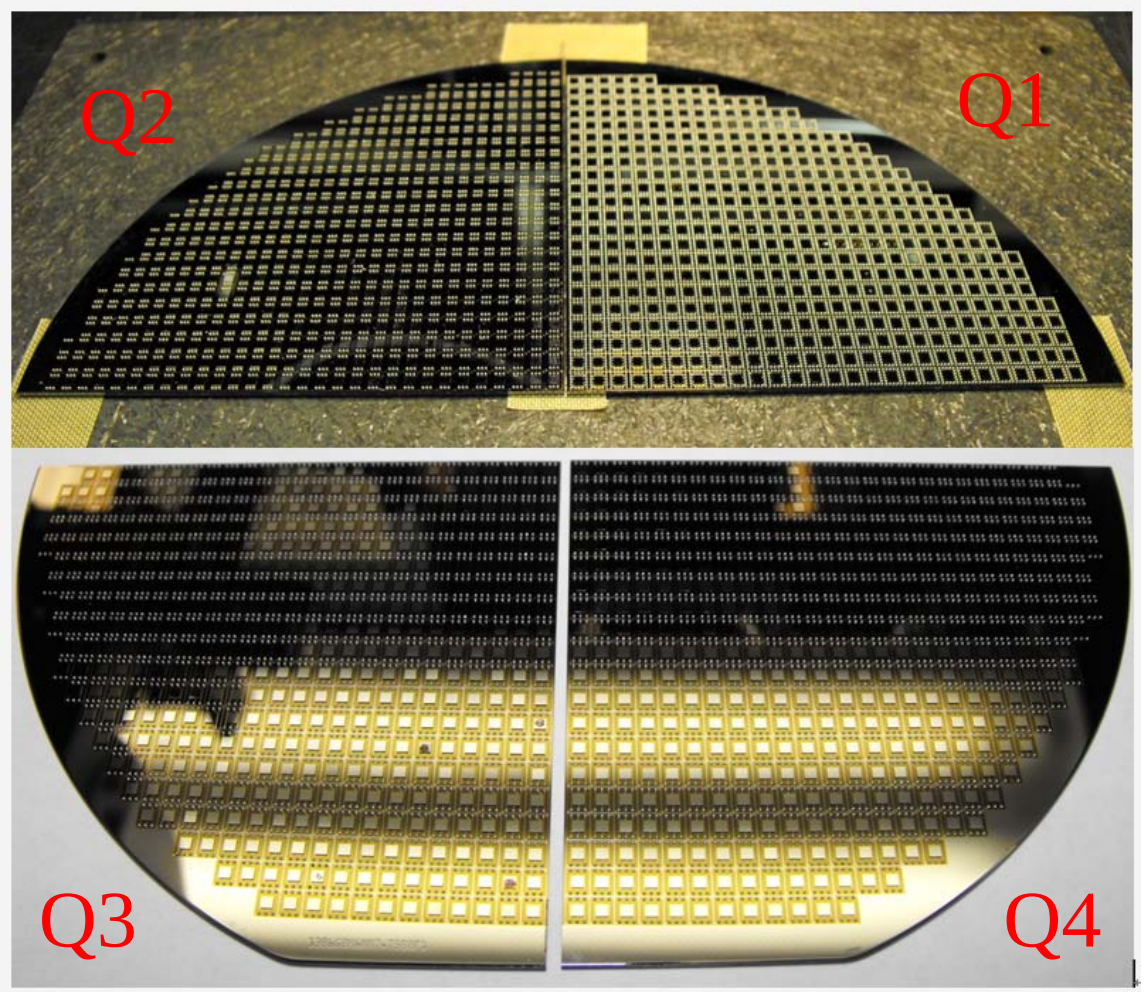


Figure 6-3 Initial Assembly on Quadrant Wafers

The initial yield on each quadrant wafer was calculated based on the probing of resistance across the daisy chain on each individual first-level package. The initial yield is shown in Table 6-3.

Table 6-3 Initial Assembly Yield of Wafer Quadrants

Quadrant	Packages Assembled	Packages Good	Yield
1	470	442	94.04%
2	506	501	99.01%
3	450	445	98.89%
4	450	446	99.11%

The initial yield was satisfactory for Quadrants 2, 3, and 4. However, Quadrant 1 had the greatest yield loss. A destructive shear test was performed on the failed Quadrant 1 packages. After removal of the flip chip die, it was found that the die was placed inaccurately as the marks (flux residue) that the solder bumps left on the CSP substrate were “shifted” from the pad location, as shown in Figure 6-4. This indicates that either the placement program or the machine vision process had a problem.

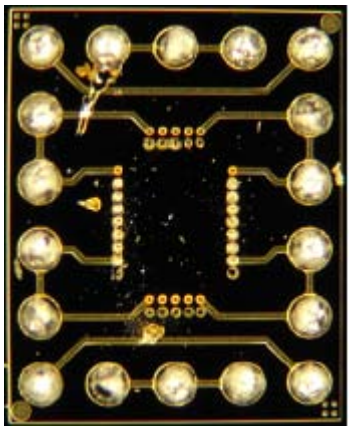


Figure 6-4 Destructive Failure Analysis of Quadrant 1 Failed Assembly

No similar failure was observed on the failed packages of Quadrants 2, 3, or 4. However, it was found from the flip chip die inspection that some bumps were either missing or damaged, as shown in Figure 6-5. Figure 6-6 shows a cracked Quadrant 2 die

captured by the die inspection. The daisy chain was broken because of the die crack. The defects on the flip chip die account for the failures on Quadrant 2, 3 4 initial assemblies.

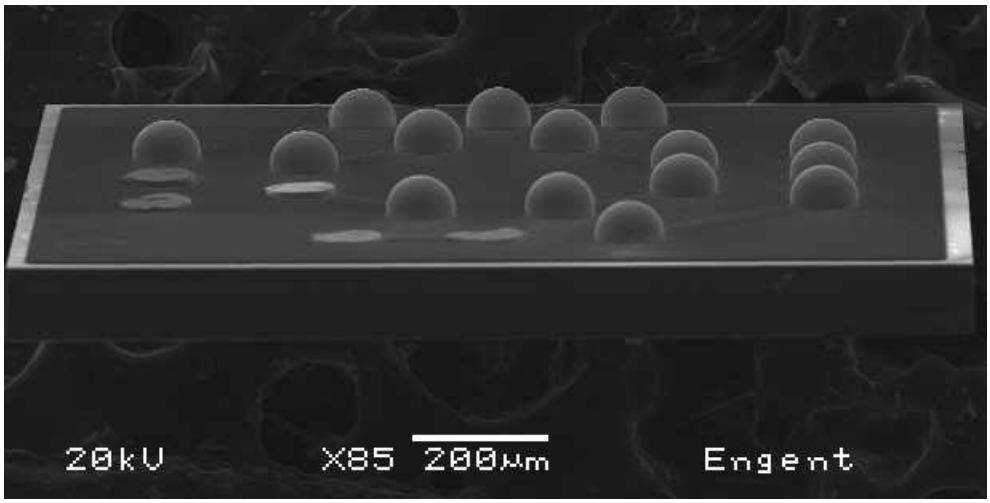


Figure 6-5 Missing Bumps on Quadrant 4 Flip Chip Die

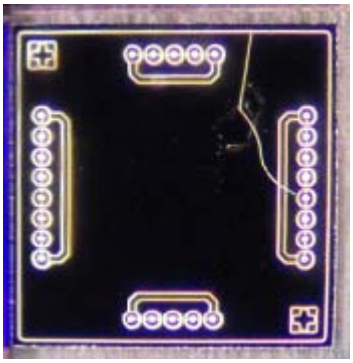


Figure 6-6 Die Crack on Quadrant 4 Flip Chip Die

6.2.4 Wafer Level Assembly Yield Improvement

Based on the failure analysis performed in Chapter 6.4, the placement error was the primary reason accounting for the yield loss on Quadrant 1 wafer. The cause of the placement inaccuracy is the machine vision of local fiducials on the first column while

performing the “step and repeat”. To solve this problem, three global fiducials on the Quadrant 1 wafer instead of local fiducials on the first column were used in the modified program.

Careful die inspection was performed on Quadrants 1, 2, 3, and 4 flip chip dies. The dies that had the bump non-uniformity issue or other defects were picked out from the waffle packs. The assembly yields on different wafer quadrants are shown in Table 6-4. 100% yield was achieved on all the quadrant wafers. The image of the assembled whole wafer that had 100% assembly is shown in Figure 6-7.

Table 6-4 Assembly Yield on Quadrant Wafers After Process Improvement

Quadrant	Packages Assembled	Packages Good	Yield
1	470	470	100.00%
2	506	506	100.00%
3	450	450	100.00%
4	450	450	100.00%

Table 6-5 shows the recorded assembly time (pick and placement) for each quadrant wafer. Considering the reflow process of 6 minutes, on average the process time for each individual 3D-WLCSP first level assembly can be calculated as follows:

$$(57+63+52+52+6) \times 60 / 1876 = 7.35 \text{ second/package}$$

Compared with the standard flip chip on board assembly process. The chip placement time plus the reflow time would be around 6-7 minutes. Thus, it can be seen that the 3D die-to-wafer assembly process with 3D-WLCSP technology can dramatically improve the throughput to around 50 times or higher.

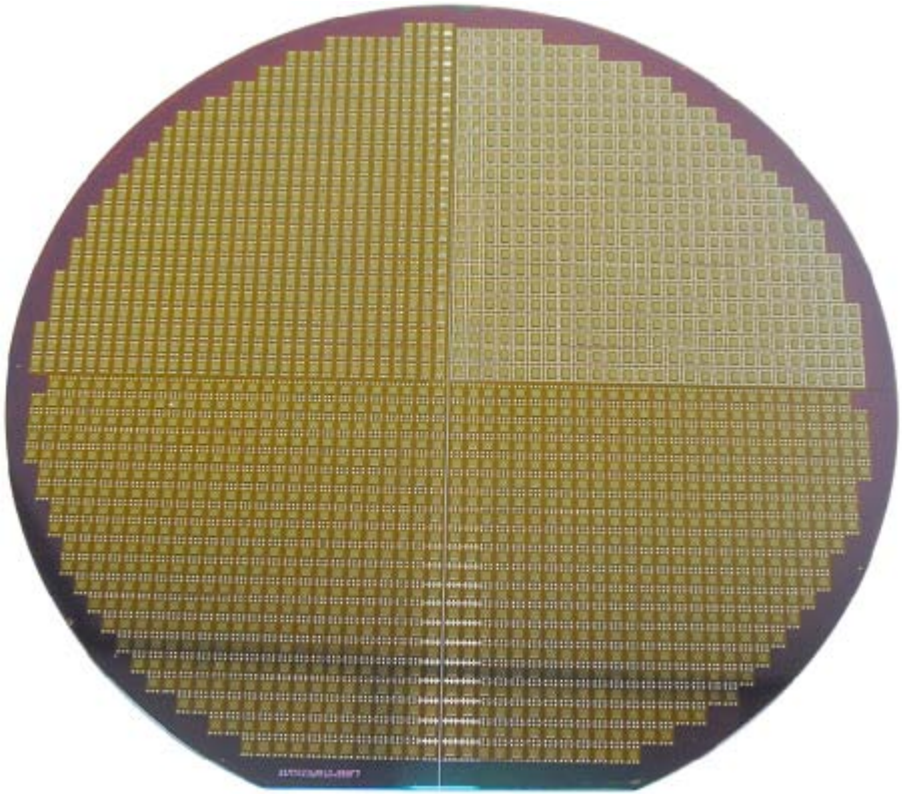


Figure 6-7 Assembled 3D-WLCSP Wafer Showing 100% Yield

Table 6-5 Assembly Time Study on Quadrant Wafers

Quadrant	Packages Assembled	Assembly Time (Minutes)
1	470	57
2	506	63
3	450	52
4	450	52

6.3 Summary

A wafer level scale-up study was carried out in this chapter. Development effort was focused on the pick and placement program creation for the wafer level assembly, flux dwell time study and the wafer level assembly yield analysis and improvement.

A “step and repeat” program technique was utilized for the programming on quadrant wafers. This technique greatly reduced the data entry and increased the machine operation efficiency.

The flux dwell time study showed that the test vehicle tiles can maintain 100% yield with a dwell time of as long as 5 hours before entering the reflow oven. This study confirmed that the flux used for the wafer level assembly will not lose fluxing functionality throughout the wafer level assembly process.

The initial assembly yield loss on Quadrant 1 was associated with the placement program vision alignment. A program improvement was performed on the Quadrant 1 pick-and-placement program. A problem resulting from bump non-uniformity was identified in the Quadrant 2, 3 and 4 flip chip test vehicles. An effort was made to remove the die with non-uniformity issue. After these efforts, all four quadrants of the test vehicle wafer showed a very robust yield of or close to 100%. The full speed of the ESEC Micron 2 machine was utilized in the wafer assembly which ensures the high throughput of the 3D-WLCSP first level assembly.

This part of the work demonstrated that the 3D-WLCSP can be assembled with high yield and thus is a proven good solution for the 3D die-to-wafer integration methodology.

CHAPTER 7 3D-WLCSP SECOND-LEVEL ASSEMBLY PROCESS DEVELOPMENT AND RELIABILITY EVALUATION

7.1 Introduction

Previous chapters have presented the assembly process development and reliability performance evaluations on the 3D-WLCSP first level packages. In this chapter, the second-level assembly, which is the assembly of first-level packaged component to FR4 organic substrate (PCB), is carried out.

In the first-level assembly, all flip chip dies were assembled face-to-face to the WLCSP wafer, as shown in Figure 6-10. After the wafer level assembly, underfill was dispensed at each individual first-level package on the wafer scale. The close-up view of a part of underfilled wafer is shown in Figure 7-1. The underfill was jet dispensed with the underfill 4 that demonstrated high reliability performance as discussed in Chapter 4. The first-level packaged wafer was then singulated by the dicing saw after the first-level underfill. The singulated first-level packages were waffle packed to get ready for the second-level assembly, as shown in Figure 7-2.

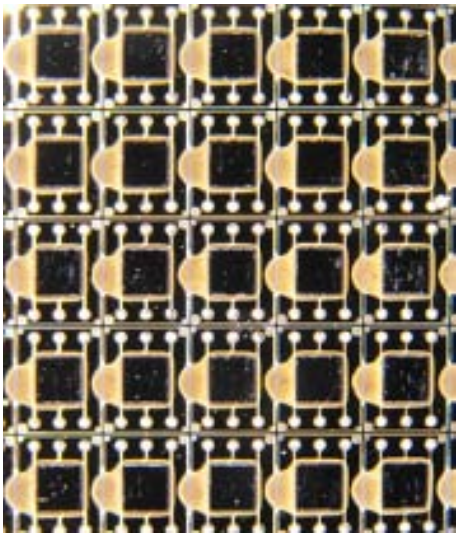


Figure 7-1 3D-WLCSP First-Level Packages After Wafer Level Underfill Dispense

Figure 7-2 Waffle Packed First-Level Packages after Wafer Dicing

A close-up view of the Quadrant 2 singulated first-level package is shown in Figure 7-3. The outline of the Quadrant 2 CSP component is 2.27mm x 3.15mm. The solder balls on Quadrant 2 WLCSP substrate have a pitch of 0.4mm.

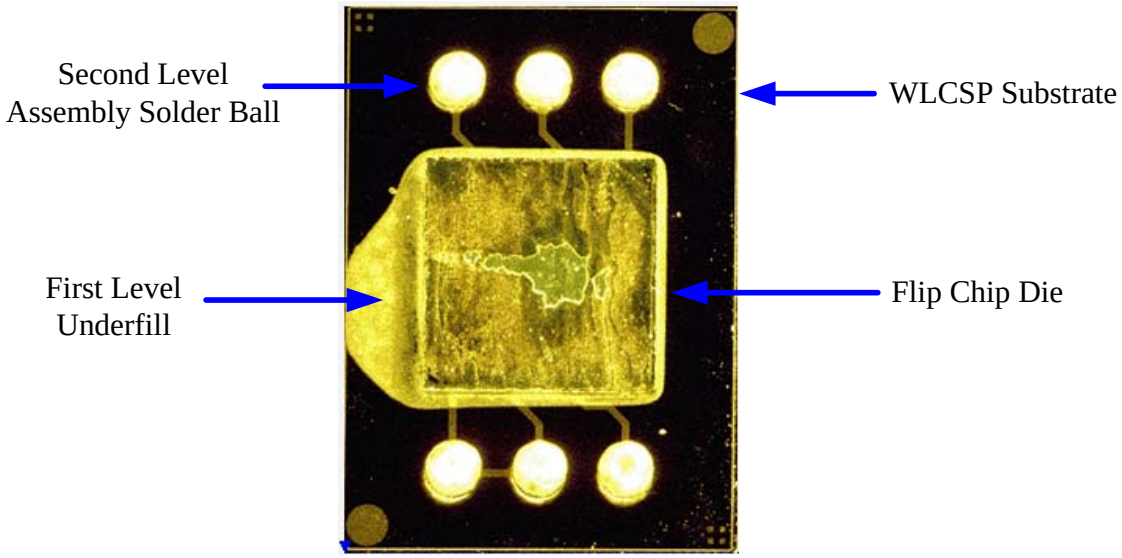


Figure 7-3 3D-WLCSP Singulated Quadrant 2 First-level Package

A standard Pb-free CSP Surface Mount Technology (SMT) assembly process was used for the second-level assembly of the 3D-WLCSP technology. Unlike the flip chip process in which the flux was applied by either dipping the component in a flux tray or pre-applied to the substrate pad by flux spray/print, the standard CSP process does not require the flux application. Instead, the solder paste is screen printed onto the substrate pads. The process flow is depicted in Figure 7-4. First the Pb-free solder paste was screen printed on the solder pads of PCB substrate. The first-level packaged CSP components were then picked and placed onto the PCB by an ESEC Micron 2 pick-and-placement machine. The assembled boards were reflowed in the Rehm RN38 convection reflow oven to form the second-level solder joints before second-level underfill was applied and cured.

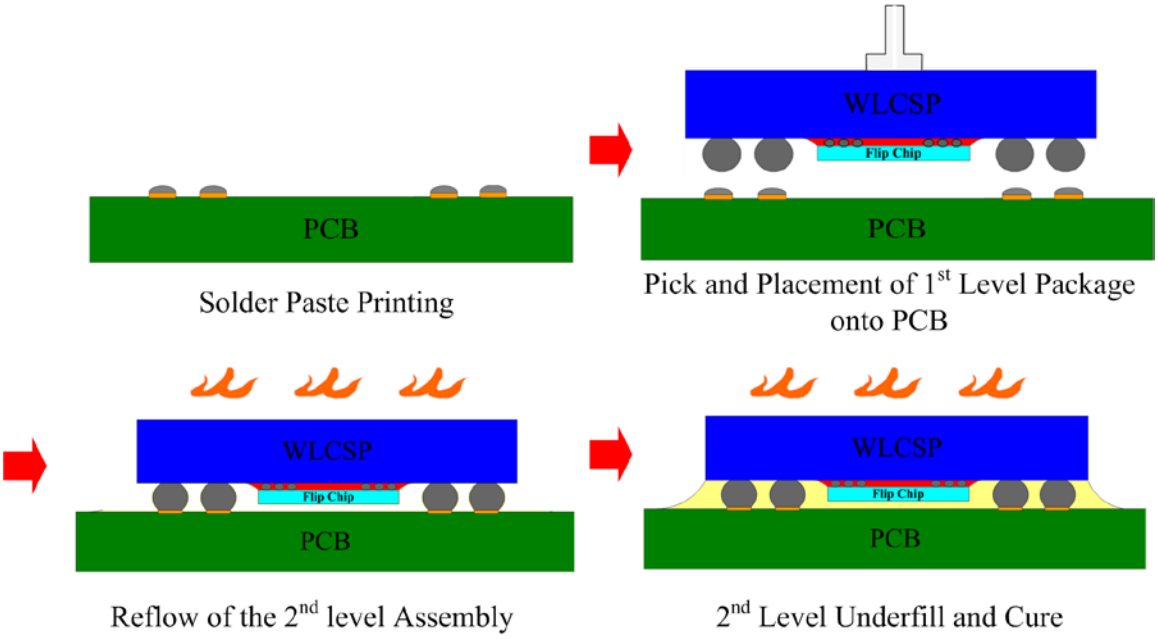


Figure 7-4 3D-WLCSP Second-Level Assembly Process Flow

7.2 Experimental Approach and Results

7.2.1 Initial Assembly Evaluation

Figure 7-5 shows a typical Pb-free reflow profile used to attach the CSP components to PCB substrate. Test vehicles Quadrant 1 and Quadrant 3 were used in this study. Quadrant 1 represents the application that has a finer pitch with CSP solder balls at the peripheral area on all four edges. The Quadrant 3 represents the application that has a coarser pitch with the CSP balls on two short edges of the peripheral area.

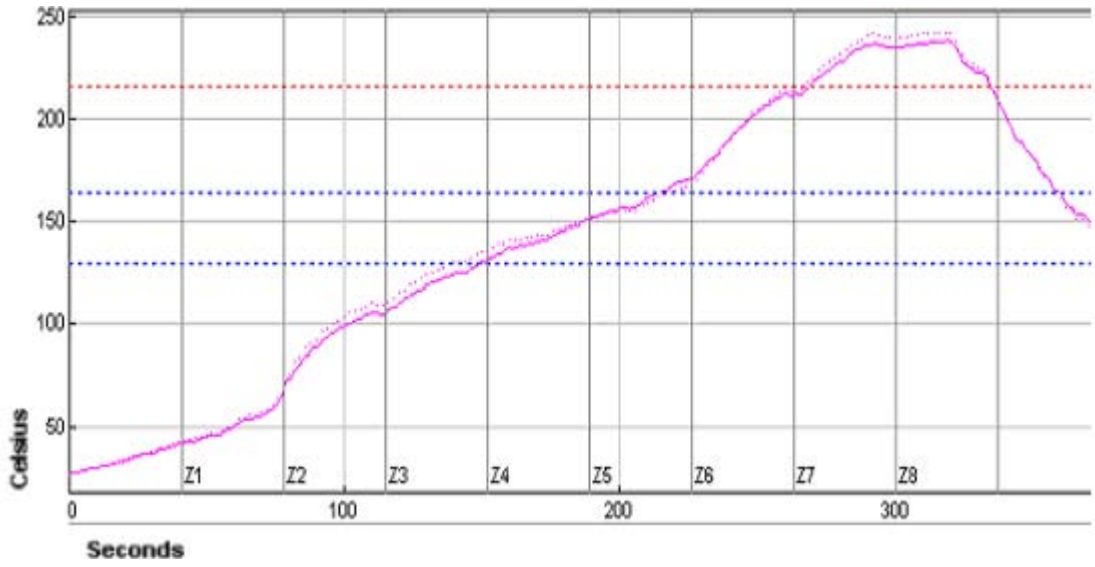


Figure 7-5 Example Reflow Profile for Pb-free CSP Assembly

Figure 7-6 shows the initial Quadrant 1 second-level assemblies on test board (PCB) after second-level capillary underfill dispense. The second-level assembly, which is a CSP to PCB test board assembly, also features a daisy chain structure that allows the probing of insulation resistance across the daisy chain as well as monitoring the package electrical continuity during the reliability testing. The daisy chain structure on the board was designed so that by probing on combination of pads, the approximate location of the failure could be easily identified. Through the probing test, whether the failure happens at the first-level interconnect (flip chip solder joints) or second-level interconnect (CSP solder joints) can be determined.

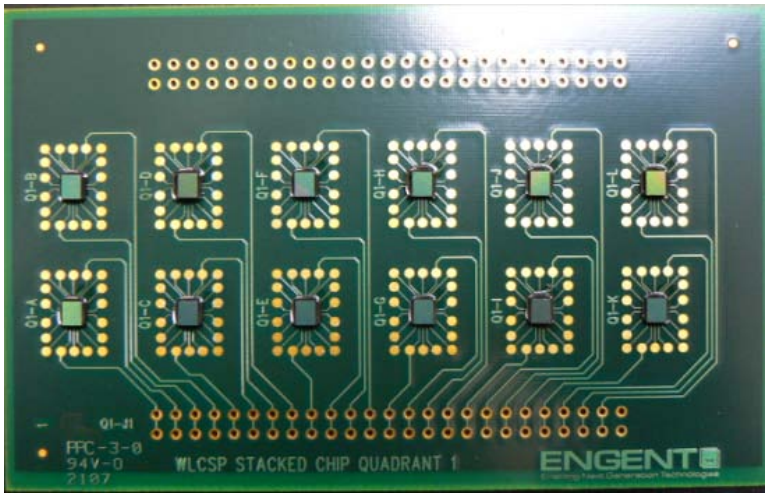


Figure 7-6 Q1 CSP on PCB 2nd Level Assembly Test Board

The initial yield on Quadrant 1 second-level assembly was 75%, according to the insulation resistance probe result across the daisy chain on the test board. Figure 7-7 shows the cross section view of the second-level CSP solder joints. It can be seen that the CSP solder ball did not mix and collapse well with the solder paste.

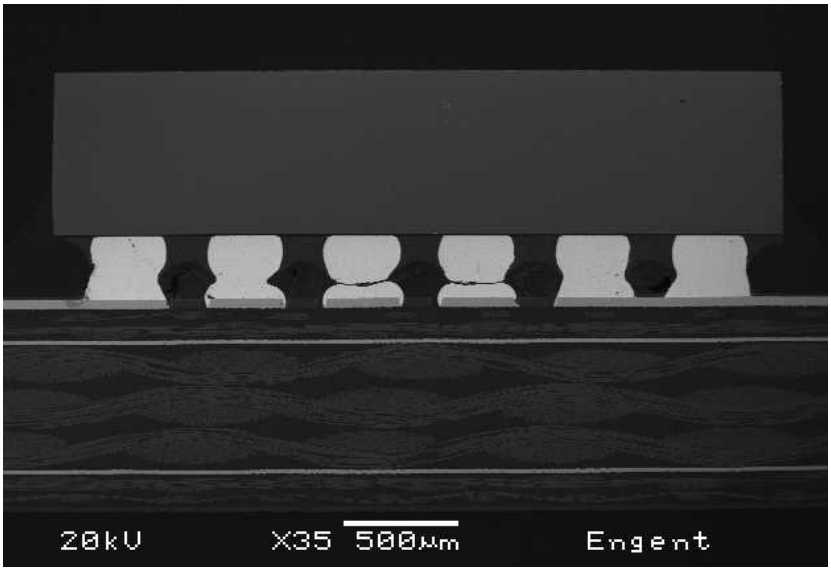


Figure 7-7 Cross Section of Initial Quadrant 1 Second Level Assembly

As discussed in previous chapters, the Quadrant 1 PA has a perimeter bumped structure with flip chip component in close proximity to the PA solder spheres. This structure creates difficulty for the first-level capillary underfill application on Quadrant 1 PA as the underfill dispensed tends to flow onto the CSP balls. A close-up view of the Quadrant 1 first level package that had the first level underfill encroachment issue is shown in Figure 7-8. The excessive underfill material acted as a barrier preventing the solder balls from adequate wetting during the second-level assembly and thus induced the yield loss.

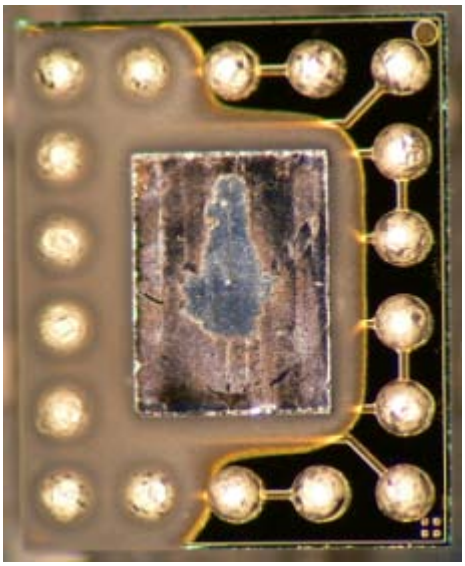


Figure 7-8 Quadrant 1 First-Level Assembly Underfill Encroachment

The other issue found in the second-level assembly is the voids in the second-level CSP solder joints, as shown in Figure 7-9. Too much voids may affect the solder joint quality, especially if the voids are formed on the solder wetting area (pad), as the voids take the place where the intermetallic compound should be formed. In this case, the solder joint integrity and the solder joint strength are undermined.

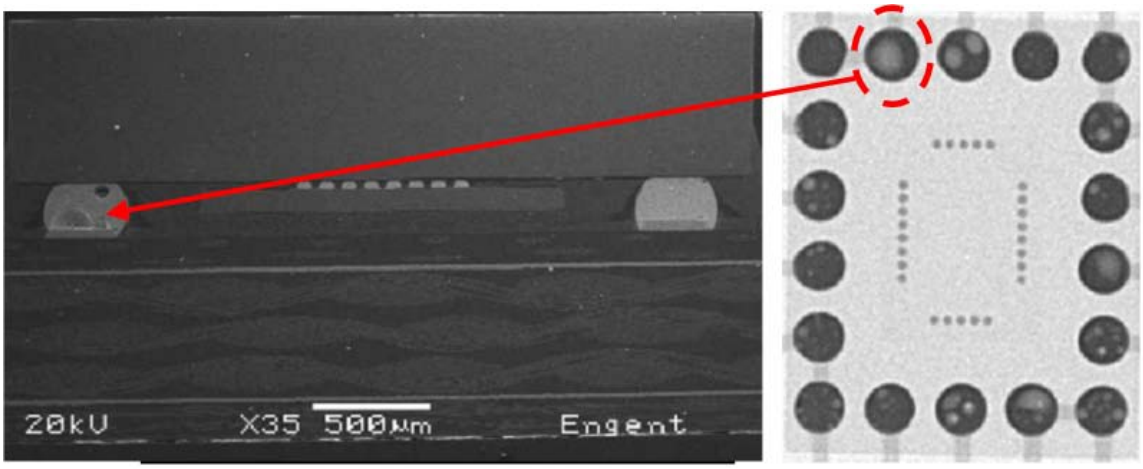


Figure 7-9 Cross Section (left) and X-Ray (right) of a Quadrant 1 Second-Level Assembly Showing Solder Joint Voids

7.2.2 Second-Level Assembly Process Improvement

Two issues were identified in the initial second-level assembly, as discussed previously. The first issue was the Quadrant 1 first-level underfill encroachment on the CSP solder balls. The initial assembly used the 25-gauge dispensing needle for the underfill application. To solve the underfill encroachment issue, a 32-gauge needle was used for its better control over the dispensed volume. Figure 7-10 shows the Quadrant 1 first level assembly with underfill applied by the 32-gauge needle. The second solution to overcome the Quadrant 1 first-level underfill encroachment issue is to apply the underfill only once after the second-level assembly. This means no underfill is applied after first-level assembly so that no underfill encroachment could happen. After the first-level wafer assembly, the wafer is diced and the singulated first-level packages undergo the second-level assembly without first-level underfill. The PCB is then reflowed to form the

second-level interconnections. After the reflow, the underfill is applied so that both first-level and second-level gaps are filled at the same time.

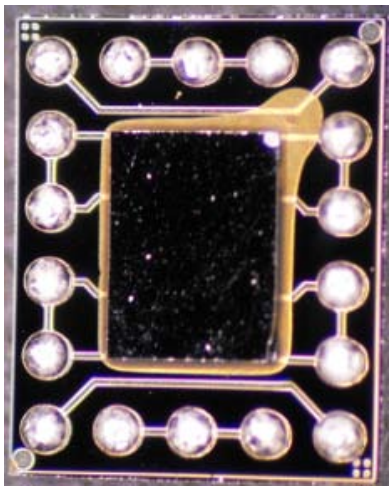
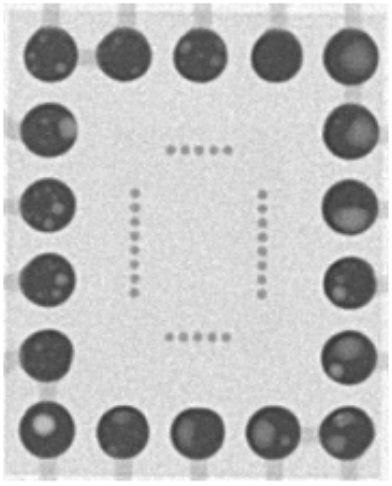


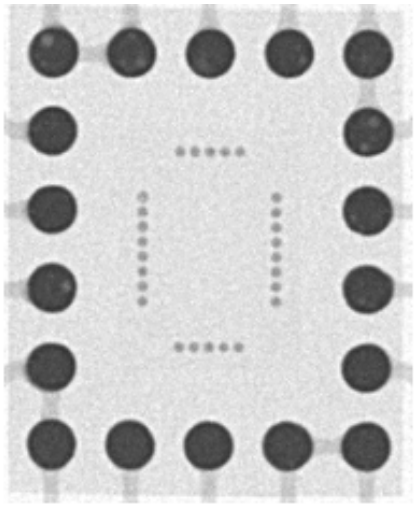
Figure 7-10 Quadrant 1 First-Level Assembly Underfill Encroachment Solution by 32 Gauge Dispensing Needle

The second issue on the initial second-level assembly was the excessive solder joint voiding in the CSP solder joints. As discussed in [39] by Ladani, longer time above liquidus and higher peak temperature will cause more solder joint voids during reflow. Based on the literature review, two different reflow profiles were designed to compare the second-level assembly solder joint voiding activities. Both Nitrogen reflow environment and air reflow environment were utilized in the two studied reflow profiles, in order to understand the impact of the reflow ambient gas on the solder joint voiding activities. The solder joint voids were captured by the X-Ray after reflow, as shown in Figure 7-11. For instance, Figure 7-11 (a) shows the sample reflowed under Nitrogen environment with 100 seconds soak time, 240 °C peak temperature and 75 seconds time above liquidus.



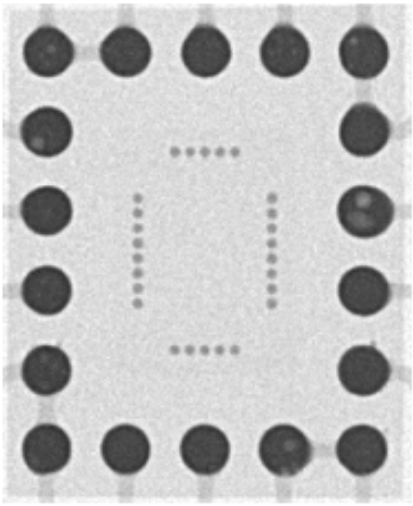
Nitrogen Reflow	
Soak Time	100s
Peak Temperature	240C
Time Above Liquidus	75s

a



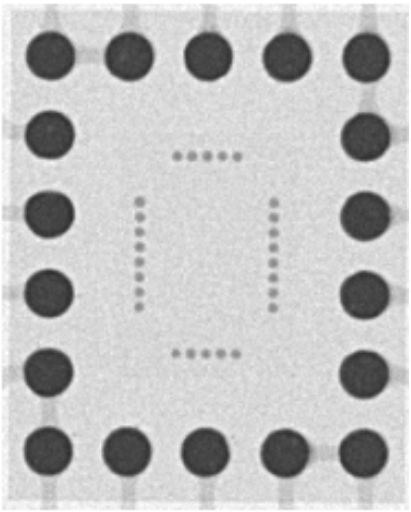
Air Reflow	
Soak Time	100s
Peak Temperature	240C
Time Above Liquidus	75s

b



Nitrogen Reflow	
Soak Time	135s
Peak Temperature	240C
Time Above Liquidus	60s

c



Air Reflow	
Soak Time	135s
Peak Temperature	240C
Time Above Liquidus	60s

d

Figure 7-11 X-Ray of the Second-Level Solder Joint

From the Figure 7-11, it can be seen that a longer soak time, lower time above liquidus will help to reduce the second-level CSP solder joint voids. Also it was found that reflow in the air environment will create fewer voids than reflow in the nitrogen environment.

100% yield was achieved in both Quadrant 1 and Quadrant 3 second-level packages with the improved reflow profile shown in Figure 7-11 (d). Figure 7-12 shows the cross section view of a Quadrant 1 second-level package assembled with the improved reflow profile demonstrating good CSP solder joint quality and minimum solder joint voids.

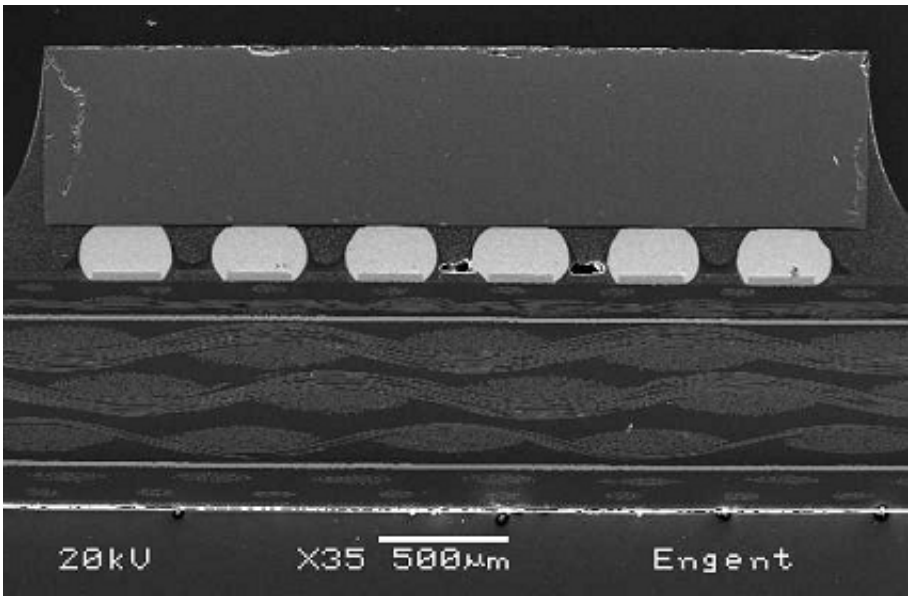


Figure 7-12 Cross Section of a Quadrant 1 Second-level Assembly

The cross section image of a Quadrant 3 second-level package is shown in Figure Figure 7-13. No underfill encroachment issue was found on Quadrant 3 second-level solder joints as the CSP solder balls were not close to the edge of the flip chip die.

.

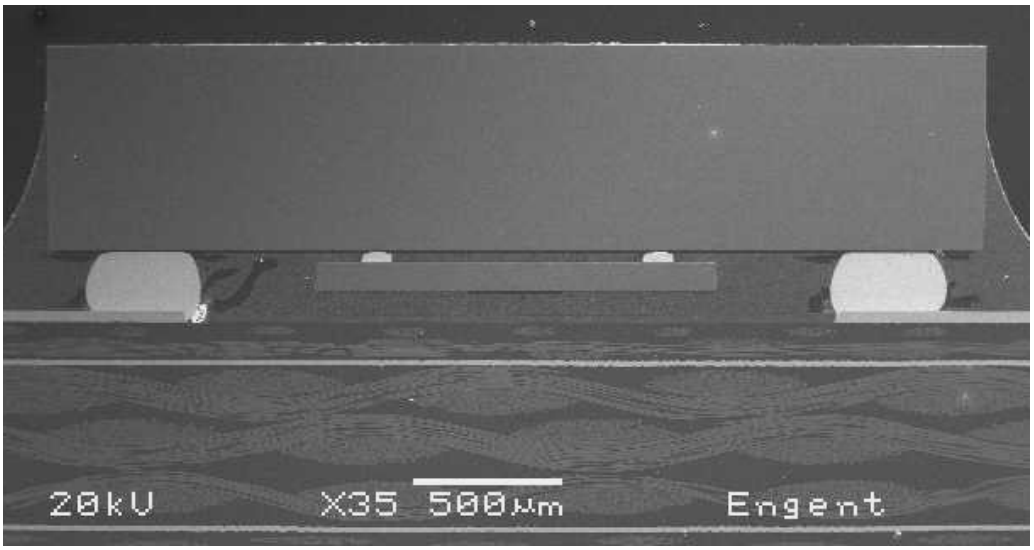


Figure 7-13 Cross Section of a Quadrant 3 Second-level Assembly

Two different solder paste materials for the Pb-free CSP assembly were used in the second-level process development. The purpose was to compare their yield and reliability performance in order to find the best solder paste candidate for 3D-WLCSP second-level assembly. The two solder paste materials differ in the flux agent and particle size. An assembly matrix containing two different solder pastes and two underfill application methods is outlined in Table 7-1, listed as assembly conditions 1 to 4. For instance, Condition 1 stands for the assembly with solder paste 1 and the underfill was applied only once after completion of second-level assembly. The single time underfill for both first and second assembly, as discussed previously, was intended to solve the

Quadrant 1 first-level underfill encroachment issue. 100% assembly yield was achieved in each condition. Condition 2 represents the assembly with solder paste 1 and the underfill was applied both after first-level assembly and second-level assembly.

Table 7-1 Quadrant 1 Second-level Assembly Matrix

Condition	Solder Paste Type	Underfill Method	Yield
1	1	One Time, After 2nd Level Assembly	18/18
2	1	After Both 1st and 2nd Level Assembly	12/12
3	2	One Time, After 2nd Level Assembly	12/12
4	2	After Both 1st and 2nd Level Assembly	12/12

For the Quadrant 3 samples, since there was no underfill encroachment issue on CSP solder balls, the underfill was applied at the end of both first-level and second-level assembly. The Quadrant 3 assembly matrix with two different solder paste materials is shown in Table 7-2. 100% yield was achieved at both assembly conditions.

Table 7-2 Quadrant 3 Second-level Assembly Matrix

Condition	Solder Paste Type	Underfill Method	Yield
2	1	After Both 1st and 2nd Level Assembly	36/36
4	2	After Both 1st and 2nd Level Assembly	24/24

7.2.3 Second-level Assembly Reliability Evaluation and Failure Analysis

All the assembled packages listed in Table 7-1 and Table 7-2 were then subject to 1500 cycles of AATC testing at a profile of cycling between –40 ºC to 125 ºC with a dwell time of 10 minutes on each side per cycle, following the JEDEC spec

JESD22-A104-B, Condition G. Electrical continuity of the daisy chain for each package was measured and recorded every 100 cycles, with a 20% change in resistance as fail/pass criteria. The purpose of this test was to find out the best assembly condition as well as the best solder paste material for the 3D-WLCSP second-level packages.

The reliability data for Quadrant 1 assemblies is reported in Table 7-3. The reliability data for Quadrant 3 packages is reported in Table 7-4.

Table 7-3 Quadrant 1 Second-level Assembly Matrix and Reliability Data

	Solder Paste/Underfill Method	AATC Cycles					
Condition		0	300	600	900	1200	1500
1	Q1 Solder 1 Underfill at 2nd level	0/18	0/18	0/18	0/18	1/18	1/18
2	Q1 Solder 1 Underfill at 1st and 2nd level	0/12	1/12	2/12	2/12	2/12	3/12
3	Q1 Solder 2 Underfill at 2nd level	0/12	1/12	1/12	1/12	2/12	2/12
4	Q1 Solder 2 Underfill at 1st and 2nd level	0/12	0/12	0/12	0/12	0/12	0/12

Table 7-4 Quadrant 3 Second-level Assembly Matrix and Reliability Data

	Solder Paste/Underfill Method	AATC Cycles					
Condition		0	300	600	900	1200	1500
2	Q3 Solder 1 Underfill at 1st and 2nd level	0/36	0/36	0/36	1/36	2/36	2/36
4	Q3 Solder 2 Underfill at 1st and 2nd level	0/24	0/24	0/24	0/24	0/24	0/24

The probe of the daisy chain on the failed Quadrant 1 packages in Conditions 1 and 2 suggest that the failures happened at the CSP solder joints rather than the flip chip solder joints. Figure 7-14 shows the cross section view of a lifted CSP solder joint from a package that was assembled with Condition 1 and failed during the reliability testing. Figure 7-15 shows the cross section view of a lifted CSP solder joint from a second-level package that was assembled with Condition 2 and failed during the reliability testing.

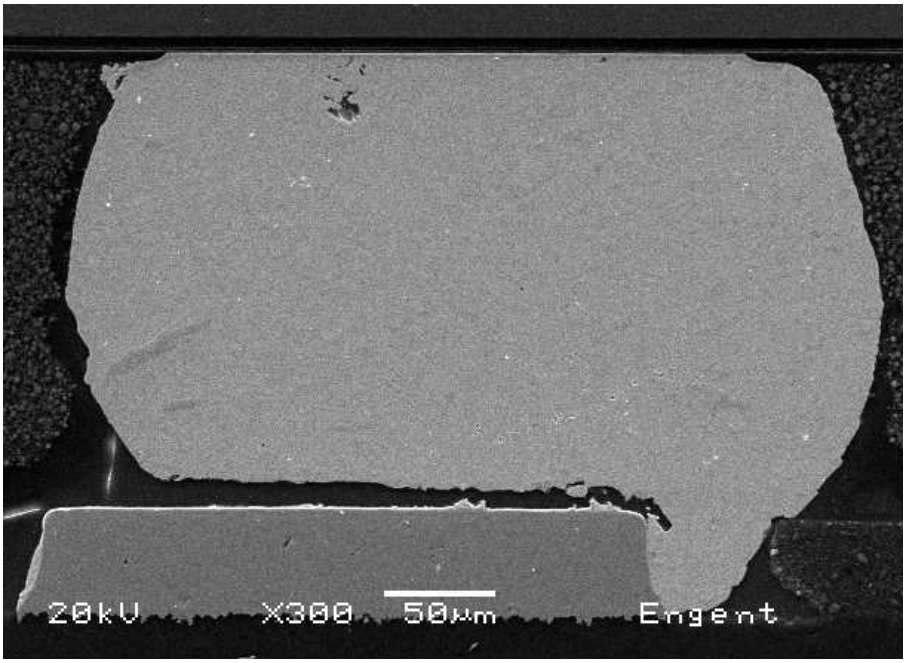


Figure 7-14 Cross Section of a Failed Quadrant 1 Second-level Package Assembled at Condition 1 after 1500 AATC Testing

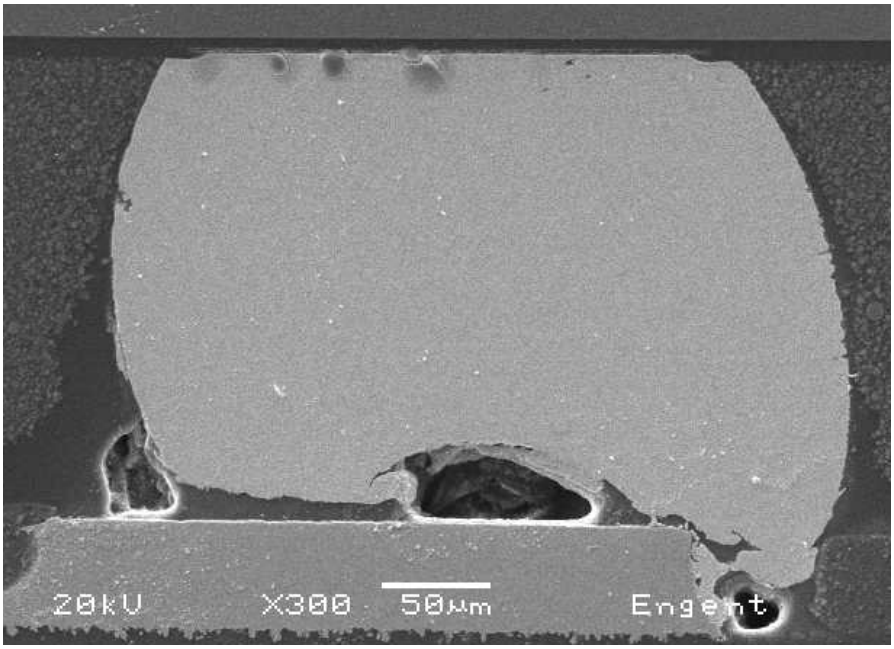


Figure 7-15 Cross Section of a Failed Quadrant 1 Second-level Package Assembled at Condition 2 After 1500 AATC Testing

By comparison, the probe of the daisy chain suggests that the failed Quadrant 1 samples in Condition 3 happened at flip chip solder joints instead of CSP solder joints. Figure 7-16 shows the cross section view of the failed flip chip solder joints in assembly Condition 3 after 1500 cycles.

Based on the reliability results shown in Table 7-3 and Table 7-4, it can be concluded that the solder paste 2 is a better candidate in terms of the reliability performance of the Quadrant 1 second-level assembly.

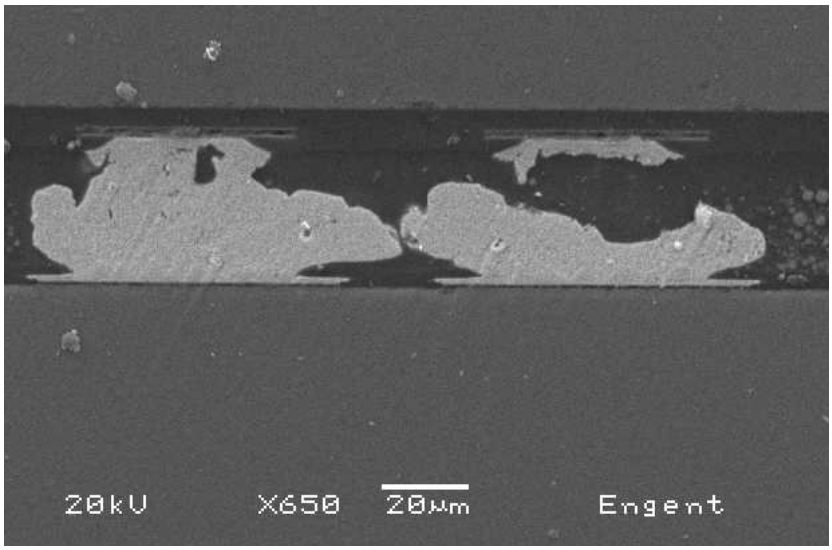


Figure 7-16 First-Level Solder Joint Failure of a Quadrant 1 Second-level Package Assembled at Condition 3

As revealed in Figure 7-16, no underfill material was observed at the surrounding area of flip chip solder joints. The reason for the absence of underfill can be explained as follows. In Condition 3, due to the application of first-level and second-level underfill at the same time, the underfill material tended to fill the big gap (200 µm) between CSP and PCB more easily and quickly than to fill the much smaller gap (30 µm) between the flip

chip and CSP. As a result, the air was trapped in the gap between the flip chip and CSP so that the voids around the first-level solder joints were formed. To solve this issue, the underfill flow speed should be reduced so that there will be enough time for the underfill to fill the gap between the flip chip and CSP at first level. This was achieved by reducing the temperature of the hot plate that carries the packages to be underfilled from 80 °C to 55 °C. The cross section of the first-level gap (between flip chip die and CSP substrate) of Quadrant 1 package assembled under Condition 3, with 55 °C hot plate temperature is shown in Figure 7-17. It can be seen that the underfill voids around first level solder joints were greatly reduced as the underfill particles filled the first-level gap well by the reduced flowing speed.

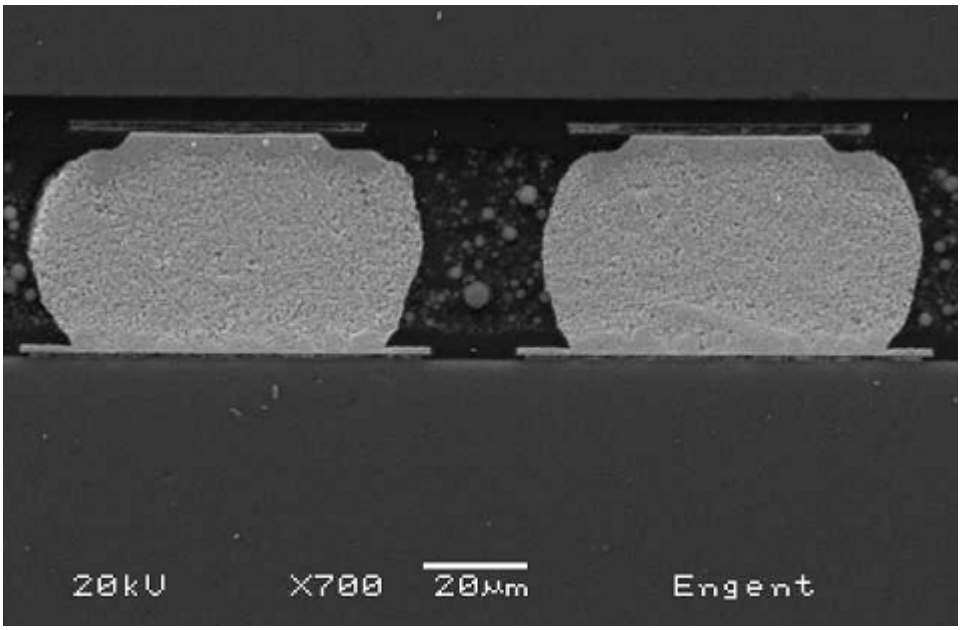


Figure 7-17 First-Level Solder Joints on Quadrant 1 Second-Level Package Showing Reduced Underfill Voids

Failure analysis was also performed on the failed Quadrant 3 second-level packages assembled with Condition 2. The daisy chain probe result also indicates that the failure happened at the second-level interconnections. Figure 7-18 shows a failed Quadrant 3 solder joint assembled with Solder Paste 1. The Quadrant 3 reliability results shown in Table 7-4 also indicate that Solder Paste 2 is a better candidate over Solder Paste 1 in terms of Quadrant 3 second-level reliability performance.

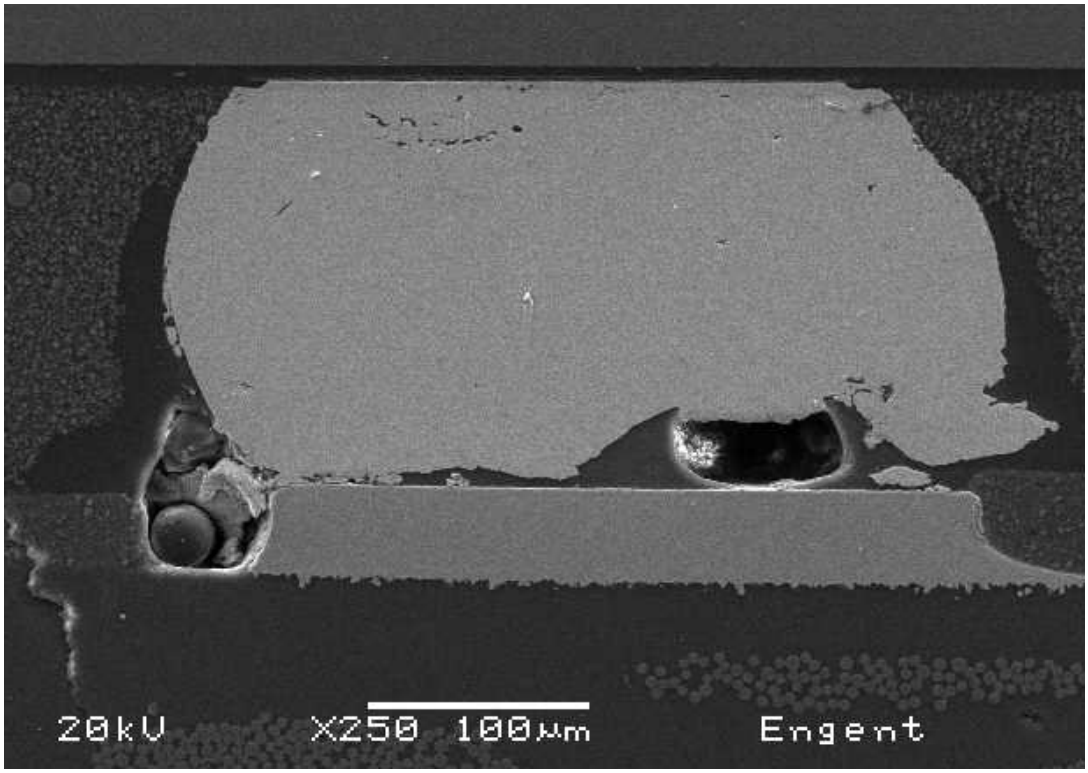


Figure 7-18 Cross Section of a Failed Quadrant 3 Second-level Package Assembled with Solder Paste 1

Figure 7-19 shows the cross section of a Quadrant 1 second-level solder joint (CSP solder joint) that was assembled with solder paste 2 after 1500 AATC testing. The

solder joint remained a good quality after the reliability testing. This indicates that the solder paste 2 is a good candidate for the 3D-WLCSP second-level assembly.

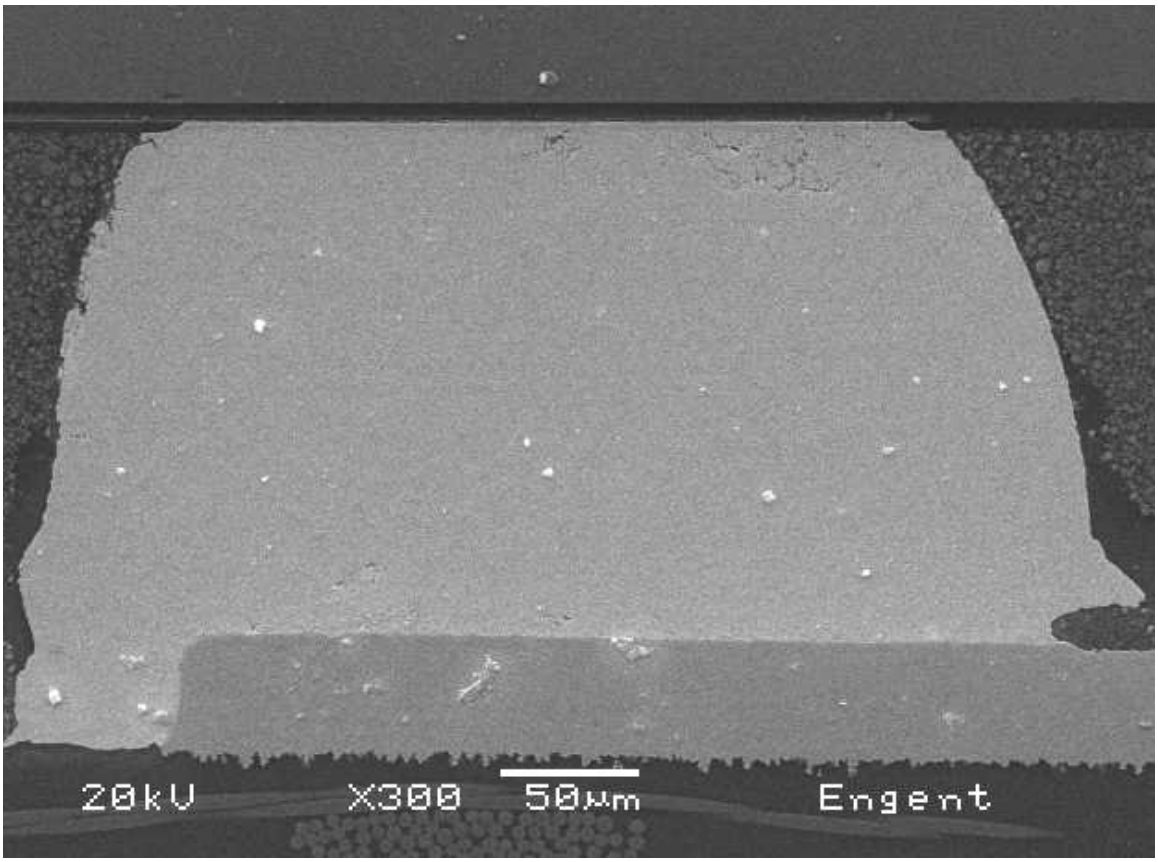


Figure 7-19 Cross Section of a Quadrant 1 Second-level Package Assembled with Solder Paste 2 after 1500 AATC Testing

7.3 Summary

The 3D-WLCSP second-level assembly shows robust yield on both finer pitch (Quadrant 1) and coarser pitch (Quadrant 3) structures. The probing across the daisy chain and the package cross sections confirmed the good quality of the solder interconnects. The reflow profile with higher soak time, lower peak temperature and time above liquidus was found to reduce voids inside CSP solder joints. In the meanwhile, the

air reflow environment was found to produce less CSP solder joint voids than the Nitrogen reflow environment.

Different underfill methods were utilized to solve the underfill encroachment issue on Quadrant 1 first-level packages. A 32 gauge dispense needle was found to have a better control over the underfill bleed-out. The option of applying underfill to both first-level and second-level packages at the same time was also studied, as an alternative to overcome the first-level underfill encroachment on Quadrant 1 packages. It was found that a lower hot plate temperature can reduce the underfill voids in the first-level package, when the underfill is applied to both first-level and second-level packages at the same time.

Two different solder paste materials were evaluated for the 3D-WLCSP second-level assembly. Solder paste 2 was identified to be a better candidate for its excellent AATC reliability performance on both Quadrant 1 and Quadrant 3 second-level packages. The reliability assessment on both first-level and second-level packages shows that the structured 3D-WLCSP packages can be manufactured with robust yield and demonstrate high thermal cycling reliability.

CHAPTER 8 CONCLUSIONS

8.1 Impact of Reflow Profile Parameters on 3D-WLCSP First-Level Assembly

The effects of the Pb-free reflow profile parameters toward the fine-pitch flip chip on silicon assembly were studied for the 3D-WLCSP first-level packages. Three reflow parameters, including soak time, peak temperature, and time above liquidus, were investigated.

It was found that 100% yield can be achieved within a wide reflow processing window if the flip chip is reflowed in a Nitrogen environment. However, the flip chip assembly yield was found to be sensitive to the reflow profile parameters if the package is reflowed in an air environment. It was found that there is a negative-correlation between package yield and flux burn time, which is the time span from the start of the soak zone to the start of the solder melting temperature, if the package is reflowed in an air environment.

None of the three reflow profile parameters studied was found to be significant in terms of the package shear strength. However, a positive-correlation was found between the package shear strength and the reflow profile ramp down rate.

The average thickness of the intermetallic compound was measured on samples built with each of the fifteen reflow profiles. It was found that the time above liquidus is the most significant factor on the average IMC thickness.

Based on a response surface methodology DOE design, LLTS reliability testing was conducted on samples assembled via fifteen different reflow profiles. Only one failure was observed on 480 tested samples after 2500 cycles. This indicated that the silicon-on-silicon package without underfill has a very robust thermal shock reliability performance.

8.2 Capillary Underfill and No-flow Underfill Flip Chip Process Development

Both a capillary flow underfill and no-flow underfill flip chip assembly process were developed for the 3D-WLCSP first-level packages.

For the capillary flow underfill flip chip assembly process, three flux materials and four underfill materials were studied on two different 3D-WLCSP test vehicles. This was done to evaluate the impact of material selections and layouts on the first-level package assembly yield and reliability. The reliability evaluation of 24 combinations of flux/underfill/layout revealed that underfill selection is the most significant factor for the 3D-WLCSP first-level assembly reliability performance. The two best-performing underfill materials were found to have the CTE values close to the CTE value of Pb-free solder. The package layout also has a significant impact on the package LLTS reliability performance. The test vehicle that has a coarser pitch and larger bumps showed better reliability performance. The flux selection was found to be non-significant for the package LLTS reliability performance.

A comprehensive first-level package qualification testing plan was carried out with the two best-performing underfill candidates. Quadrant 3 test vehicles showed only one failure up to 5000 cycles of AATC testing when the material #3 was applied by the

needle dispensing method. No failure was observed on Quadrant 3 packages up to 4000 AATC cycles when the material #4 was applied by the jet dispensing method. Quadrant 2 packages with 85-micron pitch show satisfactory reliability performance with a Weibull life of 2442 cycles. Solder joint fatigue induced by thermal cycling was the root cause for Quadrant 2 and Quadrant 3 failures. Minimal underfill delamination was observed in the assemblies throughout the AATC testing.

A no-flow underfill flip chip assembly process was also developed for the 3D-WLCSP first-level packages for higher throughput and reduced processing cost. Two no-flow underfill materials were evaluated. NF2 was identified as a better candidate for its wider processing window. Underfill volume was controlled in order to form proper underfill fillet and prevent “chip floating” phenomenon. The methodology of no-flow underfill reflow profile setup and improvement were successfully developed. Both NF1 and NF2 achieved 100% yield with underfill voids reduced.

8.3 3D-WLCSP Wafer Level Assembly

A pick and place program for a test vehicle wafer with 1,876 placement locations was developed with the application of the “step and repeat” programming method. A flux dwell time study was conducted and the flux material used for the wafer level assembly was found to be qualified for five hours of dwell time without compromising the yield.

The initial wafer level assembly was evaluated and the incorrect vision alignment on local fiducials during the “step and repeat” was identified to be the root cause for the initial yield loss for Quadrant 1 wafer. Global fiducials were used instead of local fiducials to help improve the program accuracy. Solder bump non-uniformity and flip

chip die defects were found to be the root causes for the initial wafer assembly yield loss on Quadrants 2, 3, and 4 quadrant wafers. 100% wafer level assembly yield was achieved through program improvement and die pre-inspection.

8.4 3D-WLCSP Second-level Assembly

Underfill encroachment on Quadrant 1 CSP solder balls was found to be the root cause for yield loss in the initial Quadrant 1 second-level assemblies. Different methods were developed to overcome this issue. A 32 gauge dispensing needle was found to be effective in terms of reducing the underfill encroachment. The option of underfilling both first-level package and second-level package together after second-level assembly was evaluated. A lower hot plate temperature was preferred in order to reduce the voids in the first-level gap when the first-level and second-level packages are underfilled at the same time.

Different solder pastes were evaluated in the assembly of second-level packages. Solder paste 2 was identified to have excellent AATC reliability performance on both finer pitch (Quadrant 1) and coarser pitch (Quadrant 3) second-level packages.

Work included in this dissertation has demonstrated that the novel 3D-WLCSP can be manufactured with robust yield and high reliability performance. The 3D-WLCSP technology is qualified to be a high throughput, low cost 3D die-to-wafer packaging solution.

CHAPTER 9 CONTRIBUTIONS AND SUGGESTED FUTURE WORKS

9.1 Contributions

The following contributions have been accomplished by conducting the research on the 3D-WLCSP manufacturing technology.

1. Revealed the impact of Pb-free reflow profile parameters toward flip chip silicon-on-silicon package yield, intermetallic compound thickness, shear strength and thermal shock reliability performance.
2. Investigated the impact of reflow profile ambient gas environment, flux amount and flux burn time (from the start of soak zone to the start of reflow zone) on the yield of silicon-on-silicon flip chip assembly.
3. Identified the flux/underfill candidates for fine pitch, thin profile, low stand-off silicon-on-silicon flip chip assembly.
4. Developed a high yield, stable process for flip chip silicon-on-silicon assembly with no-flow underfill assembly process.
5. Developed the high yield (100%), high throughput, 3D die-to-wafer flip chip assembly technique.
6. Developed the Pb-free CSP assembly process parameters on solder joint quality improvement and voids reduction.

9.2 Suggested Future Works

Both the capillary underfill assembly process and the no-flow underfill assembly process have been developed for 3D-WLCSP first-level packages to achieve the 3D die-to-wafer integration. The wafer level applied underfill assembly technique, which has the underfill material pre-applied to the flip chip wafer, can be developed in the future for even higher throughput for the first-level assembly.

The study has shown that the “flux burn time” is critical to the silicon-on-silicon flip chip assembly yield when the reflow is performed in the air environment. The flux material is consumed in two ways. The first way is the chemical reaction in which the flux removes the metal oxidation on the surface of solder and pad. The second way is the evaporation. A deep study over these two phenomena can be performed so that the package yield could be estimated based on the flux amount and reflow process parameters.

With the reliability data and failure analysis on both first-level and second-level packages with different processing materials such as underill, the finite element analysis (FEA) modeling can be performed for 3D-WLCSP package simulation and prediction.

The through silicon vias (TSVs) is getting the industry attention. The TSVs can be fabricated through the backside of the WLCSP wafer so that higher level of packaging integration and further packaging densities based on the 3D-WLCSP technology can be achieved.

REFERENCES

[1] C.A. Harper, D.F. Baldwin, et al. Electronic Packaging And Interconnection Handbook, 4th Edition, New York, McGraw-Hill, TK7870.15.H38, 2005.

[2] Y. Liu, “Lead-Free Assembly and Reliability of Chip Scale Packages and 01005 components”, PhD Dissertation, 2006, Auburn University.

[3] L. F. Miller, “Controlled Collapse Reflow Chip Jointing”, IBM Journal of Research and Development, vol. 13, No. 3, 1969.

[4] Amkor Flip Chip Packaging Technology Solution Sheet, http://amkor.co.kr/datasheets/flipchip.pdf

[5] J. H. Lau, Flip Chip Technologies: McGraw-Hill, 1990.

[6] R. Tummala, Fundamentals of Microsystems Packaging. New York: McGraw Hill, 2001.

[7] J. Giesler, G. O. Malley, M. Williams, and S. Machuga, "Flip chip on board connection technology: process characterization and reliability," IEEE Transactions on Components, Packaging, and Manufacturing Technology, vol. 36, pp. 449-450, 1996.

[8] M. Abtew, G.Selvaduray, “Lead-free Solders in Microelectronics,” Materials Science and Engineering: R: Reports, Vol. 27, Issues 5-6, pp. 95-141, 2000.

[9] V. Wakharkar, C. Matayabas, S. E. Lehman, R. Manepalli, M. Renavikar, S.Jayaraman and V. LeBonheur, "Materials Technologies for Thermomechanical Management of Organic Packages," Intel Technology Journal, vol. 9, November, 2005.

[10] S. Michaelides and S. K. Sitaraman, "Role of Underfilling Imperfections on Flip-Chip Reliability," ASME International, InterPack, vol. 2, pp. 1487-1493, 1997.

[11] T. Soga, “Semiconductor Resin Package Structure,” United States Patent: 4825284,1989.

[12] G.R. Blackwell, “The Electronic Packaging Handbook,” CRC and IEEE press, Chapter 4, 2000.

[13] I.J. Rasiah, “Selecting an Underfill for Flip Chip Packaging,” Electronics Engineer, May 1999.

[14] K. Gilleo, “The Chemistry & Physics of Underfill”, NEPCON West 1998, No. 1, pp. 280-292, 1998.

[15] D. Lu and C.P. Wong, Materials for Advanced Packaging, Springer Science +Business Media, LLC, pp. 315, 2009.

[16] D. W. Milner, "Application assessment of high throughput flip chip assembly for a high lead-eutectic solder cap interconnect system using no-flow underfill materials," IEEE Transactions on Electronics Packaging Manufacturing, vol. 24, p. 307, 2001.

[17] N.W. Pascarella, D.F. Baldwin, “Compression flow modeling of underfill encapsulants for low cost flip-chip assembly”, IEEE Transactions on

Components, Packaging, and Manufacturing Technology-Part C, vol.21, No.4, October 1998.

[18] S. Lee, R. Master, and D. F. Baldwin, "Assembly Yields Characterization of High I/O Density, Fine-pitch Flip Chip in Package using No-Flow Underfill," Electronic Components and Technology Conference, 2007.

[19] S. Lee, R. Master, and D. F. Baldwin, "Assembly Yields Characterization and Failure Analysis of Flip Chip in Package Using No-Flow Underfill," International Wafer Level Packaging Congress, pp. 169-175 2007.

[20] S. Lee, M. J. Yim, R. Master, C.P.Wong, and D. F. Baldwin, "Void Formation Study of Flip Chip in Package Using No-Flow Underfill," IEEE Transactions on Electronics Packaging Manufacturing, 2007.

[21] B.A. Smith, R. Thorpe, D.F. Baldwin, “A process and reliability analysis of no-flow underfill materials for high throughput flip chip processing,” Electronics Manufacturing Technology Symposium, Twenty-Sixth IEEE/CPMT International, pp.178 – 190, 2000.

[22] S. Lee, D.F. Baldwin, “Near Void-Free Assembly Development of Flip Chip Using No-Flow Underfill,” IEEE Transactions on Electronics Packaging Manufacturing, vol. 32, No. 2, April 2009.

[23] “Chip Scale Packaging (CSP) Technology,” http://extra.ivf.se/ngl/documents/ChapterD/ChapterD1.pdf.

[24] "Wafer-Level Packaging," http://www.siliconfareast.com/wl_package.htm.

[25] R. Chilukuri, D. Hays, “Wafer Level Packaging: Yesterday, Today and Tomorrow,” Future Fab Intl. Issue 21, 2006.

[26] P. Garrou, “Wafer level chip scale packaging (WL-CSP): an overview”, IEEE Transactions on Advanced Packaging, Volume 23, Issue 2, pp. 198-205, 2000.

[27] J. H. Lau, “Critical Issues of Wafer Level Chip Scale Package (WLCSP) With Emphasis on Cost Analysis and Solder Joint Reliability,” IEEE Transactions on Electronics Packaging Manufacturing, Vol. 25, No. 1, pp.42-50, 2002.

[28] J. Perkins, "Overview of Packaging Technologies & Markets," International Wafer-Level Packaging Conference, Oct. 2009.

[29] M. Kada, L. Smith, "Advancements in Stacked Chip Scale Packaging (S-CSP), Provides System-in-a-Package Functionality for Wireless and Handheld Applications,"Pan Pacific Microelectronics Symposium Conference, Jan.2000.

[30] M. Karnezos, “3D packaging: where all technologies come together”, Electronics Manufacturing Technology Symposium, 2004. IEEE/CPMT/SEMI 29th International.Jul 14-16, pp. 64 – 67, 2004.

[31] http://www.evgroup.com/en/solutions/logic_memory/c2w_bonding/

[32] T. G. Tessier, D. Scott, D. McComb and R. Forcier, “Mobile Handsets Dictate Flip Chip and Wafer Level Packaging Trends”, Semicon West, July 2008.

[33] S. Barrett, J. Reche, D-H Kim and D. Stepniak, “A New Wafer Level Package for Improved Electrical and Reliability Performance”, 28th Electronics Manufacturing Technology Symposium, IEMT (2003), pp. 157 – 169, July 2003.

[34] P. Elenius, **“**Flex on Cap-Solder Paste Bumping”
47th Electronic Components and Technology Conference, 1997.

[35] N. Lee, "Optimizing the reflow profile via defect mechanism Analysis," Soldering & Surface Mount Technology, pp. 13-20, 1999.

[36] M.R. Harrison, “Lead-free reflow soldering for electronics assembly”, Soldering & Surface Mount Technology, 13/3, pp. 21-38, 2001.

[37] B. Salam, C. Virsead, H. Da, N. N. Ekere, and R. Durairaj, "Reflow profile study of the Sn-Ag-Cu solder " Soldering & Surface Mount Technology, vol. 16/1, pp. 27-34, 2004.

[38] J. Pan, "The effect of reflow profile on SnPb and SAC solder joint shear strength," Soldering & Surface Mount Technology, vol. 18/4, pp. 48-56, 2006.

[39] L. J. Ladani, "Effect of selected process parameters on durability and defects in SMT assemblies for protable electronics," IEEE Transactions on Components and Packaging, vol. 31, pp. 51-60, 2008.

[40] P. L. Tu, Y. C. Chan, and J. K. L. Lai, "Effect of intermetallic compounds on the thermal fatigue of surface mount solder joints," IEEE transactions on Components and Packaging, vol. Part B, pp. 87-93, 1997.

[41] J. H. Lau, "Solder joint reliability of flip chip and plastic ball grid array assemblies under thermal, mechanical and vibrational conditions," IEEE Transactions on Components and Packaging, vol. Part B, pp. 728-735, 1996.

[42] V. K. Garj, "Reliability on surface mount solder joint," Proceedings of the Fourth International Microelectronics Systems, pp. 19-23, 1995.

[43] A. C. K. So and Y. C. Chan, "Reliability studies of surface mount solder joints—effect of Cu–Sn intermetallic compounds," IEEE Transactions on Components and Packaging Manufacturing Technology-Part B, vol. 19, pp. 661-668, 1996.

[44] A. C. K. So, Y. C. Chan, and J. K. L. Lai, "Aging study of Cu–Sn intermetallic compounds in annealed surface mount solder joints," IEEE Transactions on Components and Packaging Manufacturing Technology-Part B, vol. 20, pp. 161-166, 1997.

[45] W. Huang, J. M.Loman, and B. Sener, "Study of the effect of reflow time and temperature on Cu–Sn intermetallic compound layer reliability," Microelectronics Reliability, vol. 42, pp. 1229-1234, 2002.

[46] D.C. Montgomery, Design and Analysis of Experiment, 6th Edition, Arizona, JOHN WILEY&SONS, INC, ISBN 0-471-66159-7, pp.405-408, 2005.

[47] F. Zhang, M. Chen, W.T. Chian, “An investigation into the effects of flux residues on properties of underfill materials for flip chip packages,” IEEE Transactions on Components and Packaging Technologies, vol. 26, pp. 233-238, 2003.

[48] G. Baynham, D.F. Baldwin, etc., “Processing and Reliability of Flip Chip with Lead-Free Solders on Halogen-Free Microvia Substrates,” Proceedings of the 25th International Electronics Manufacturing Technology Symposium, pp. 263-269, Santa Clara, CA, October, 2000.

[49] G. Baynham, D.F. Baldwin, etc., “Flip Chip with Lead Free Solders on Halogen Free Microvia Substrates,” Proceedings of the Electronic Components and Technology Conference, pp. 1135-1139, Orlando, FL, May, 2001.

[50] R. Thorpe, D.F. Baldwin, “Yield analysis and process modeling of low cost, high throughput flip chip assembly based on no-flow underfill materials,” IEEE

Transactions on Electronics Packaging Manufacturing, vol. 24, Issue 2, pp. 123-135, 2001.

[51] J. Lu, “3D Integration: Why, What, Who, When?” FUTURE FAB International, Issue 23, 2007.