Stacked Solder Bumping Technology for Improved Solder Joint Reliability

Xingsheng Liu and Guo-Quan Lu

Power Electronics Packaging Laboratory
Center for Power Electronics Systems
Virginia Tech, Blacksburg, VA 24061-0111

Shuangyan Xu, and David A. Dillard
Adhesion Laboratory
Engineering Science and Mechanics Department
Virginia Tech, Blacksburg, VA 24061

Abstract

Reliability in flip chip and Ball Grid Array is a major concern. Solder joint geometry is extremely important from a reliability point of view. Commonly, the flip chip solder bump takes on the shape of a spherical segment. Mathematical calculations and finite element modeling have shown that the hourglass-shaped solder bumps experience the lowest plastic strain and have the longest lifetime. In an effort to optimize solder bump connections, we have developed a stacked solder bumping technique for fabricating triple-stacked hourglass-shaped solder bumps. This solder bumping technology can easily control the solder joint height and shape. The structure of triple-stacked solder bumps consists of an inner bump, middle bump and external bump. The triple-stacked solder bumps have better compliance and are able to relax the stress caused by the coefficient of thermal expansion (CTE) mismatching between the silicon chips and substrates since it has greater height. Furthermore, the hourglass-shaped solder bump has a much lower stress/strain concentration at the interface between the solder bump and the silicon die as well as the interface between solder bump and substrate than the barrel-shaped solder bump, especially around the corners of the interfaces. In this paper, the solder bumping process is elaborated and reliability is evaluated. Mechanical tests have been carried out to analyze the strain and stress of the solder joints. The interfaces of the solder joints are investigated by SEM and EDX. Experimental results show that the triple-stacked hourglass-shaped solder bump is more reliable than the traditional solder bump.

1. Introduction

One of the most promising technology of today’s chip-level interconnection is flip-chip technology, which has emerged as a low-cost, high density, and high performance interconnection method for integrated-circuit chips [1-3]. Ball-grid array (BGA) becomes the choice of the first-level and second-level interconnection as the trend toward higher input/output (I/O), higher performance, and higher yield continues [2, 4]. These area array packages (flip chip, chip scale packages (CSP) and BGA) require the formation of solder bumps for the chip and board assembly. However, solder joint fatigue is a serious concern in area array packages. Thermal stresses caused by local and global CTE mismatch are the main reliability issues for solder bump interconnections. Numerous factors affect solder bump fatigue performance, such as chip size, joint geometry, interface metallurgy, underfill and substrate. Of these factors, solder joint geometry plays an important role. The solder bumps processed by the conventional solder bumping technologies take on the shape of a spherical segment. Deviations from the spherical shape constraint can significantly increase or decrease the thermal cycle lifetime [5-7]. Several approaches have been taken to increase the standoff height and control the shape of solder bump connection [5, 8-13]. However, some of these technologies are not easily implementable in a manufacturing environment and some of them have yet to prove their cost-effectiveness. Therefore, new solder bumping technologies, which can control the solder joint geometry and shape, need to be developed.

Several processes have been demonstrated to deposit solder bumps on a wafer or chip. These include printing, plating, evaporation, sputtering, wire bumping, and dispensing. The choice of one technology over another is influenced primarily by the bump dimensions and pitch, composition, and cost. Evaporation, electroplating and stencil printing are the most widely practiced methods due to their cost effectiveness. Recent advances in stencil technology and solder pastes have boosted stencil-printing technology in flip chip applications. This paper introduces a stacked solder bumping technology based on stencil printing technique. The stacked solder bumping technology can easily control the geometry and shape of the solder joint and thus optimize solder bump connections for improved reliability. In this paper, the stacked solder bumping technology is reported for the power
semiconductor (such as insulated gate bipolar transistor (IGBT), diode) interconnection application as the original motivation of this work is to improve the solder bump lifetime for power semiconductor interconnection. However, this technology is not limited to power electronics packaging and applicable to other solder bump area array packaging.

Inside the state-of-the-art power devices and modules, interconnection of power devices is accomplished with wire bonds, which are prone to noise, parasitic oscillations, and fatigue failure. We introduced Die Dimensional Ball Grid Array (D²BGA) power chip-scale package, solder bump flip-chip technology for interconnecting power chips and developed a three-dimensional structure for packaging Integrated Power Electronics Modules (IPEMs) [14-17]. Figure 1 illustrates the structure and cross section of the D²BGA package. This package consists of a power chip, inner solder bumps, high-lead solder balls, and molding resin. The surface of the package has an array of solder bumps that are connected to the chip pads by inner solder bumps and under bump metals. It has the same lateral dimensions as the starting power chip. The proposed three-dimensional IPEM structure is shown in Figure 2. We have demonstrated the improved electrical and thermal performances of the D²BGA power chip-scale packages as well as power modules [14-15]. The reliability requirement for power electronics modules is rather demanding. For some applications, the lifetime of ~30-year is expected [18]. This paper addresses the issue of improving the power chip solder bump interconnection reliability by developing the stacked solder bumping technology as an effort to optimize solder bump geometry and shape.

In the following sections, we first discuss the solder bump geometry and shape design and materials selections, then we introduce the triple-stacked solder bumping process and flip chip bonding process. Finally the reliability assessment and failure analysis of the solder bumps are presented.

2. Solder Bump Geometry and Shape Design and Materials Selection

Thermally induced stresses are the main reliability issues for flip-chip interconnections and BGA technology. As we mentioned earlier, local CTE mismatch among the silicon chip, solder bump, underfill encapsulant, and substrate is one of the sources of thermal stresses. The other major thermal stress source is the global CTE mismatch between the silicon chip and the substrate. This global CTE mismatch causes high shear strain in the solder bump interconnection during thermal cycling. Coffin-Manson equation [19] is the widely used relationship to predict the solder bump fatigue lifetime [1-2].

\[ N_f = \theta (\Delta \gamma_p)^{\phi} \]

where \( N_f \) is the number of cycles to failure; \( \Delta \gamma_p \) is plastic strain range and \( \theta \) and \( \phi \) are material constants. For solder joints, \( \theta \) is about 1.3 and \( \phi \) is about 2 [20-21]. The strain imposed on the solder joints can be estimated by the following simplified expression,
where $\Delta \gamma$ is the shear strain imposed, $\Delta \alpha$ is the difference in the coefficient of expansion between the joined materials, $\Delta T$ is the temperature change, $a$ is the distance from the neutral expansion point of the joined materials, and $h$ is the height of the interconnect. In practice, it is not enough to use the above oversimplified expression to calculate the thermal strain. Also Coffin-Manson equation is modified to more accurately predict the solder bump lifetime. However, in modified Coffin-Manson models, the number of cycles to fail is still proportional to $(\Delta \gamma)^{\phi}$. In order to maximize the fatigue lifetime, we need to minimize the thermal strain. In a packaged module, an optimized solder joint geometry for device interconnection reduces the shear strain. As we can see from the simplified thermal strain expression, the shear strain decreases as the height of the solder bump increases. Thus, one potential approach is to use greater standoff solder interconnects with an aspect ratio much greater than one [22-23]. A tall compliant solder joint would enhance the CTE mismatch absorption capability.

Solder joint shape is also extremely important from a reliability point of view. In Figure 3, we present a few typical solder bump shapes that could be used for solder bump interconnections. Commonly, the flip chip solder bump takes on the shape of a spherical segment. Deviations from the spherical shape constraint can significantly increase (stretched joint) or decrease (squashed joint) the thermal cycle lifetime [24-25]. Mathematical calculations and finite element modeling have shown that the stress-strain behavior in the BGA solder joints is greatly affected by the shape of BGA bumps and a symmetric hourglass-shaped solder bump experiences the lowest plastic strain and has the longest lifetime [24-26]. Generally, the solder bump failure occurs first at the interfaces between solder bump and silicon chip, and solder bump and substrate. This is due to the high thermal stress concentration at these adhering interfaces, especially at the corners [1, 26, 28]. It is commonly known that, for fatigue failure, minute cracks start at one or more points that have high localized stress and gradually spread by fracture of the material at the edges of the cracks where the stress is highly concentrated. As we know, the stress and strain field near bi-material bonding or contact edges show singular behavior, which can induce a considerably larger stress than the nominal stress. Therefore, it is very important to reduce the stress singularity as much as possible as a way to improve the reliability of solder joints. Among the several common solder joints shown in Figure 3, the contact edges of the hourglass type shows the smallest singularity, compared with the other two types of solder joints, as can be qualitatively explained according to the analysis in [29-30]. It is shown in [29-30], as the contact angle increases, the singularity increases and becomes more significant. Therefore, the hourglass-shaped solder bump has a much lower stress concentration at the corners than the barrel-shaped solder bump. As a result, solder bump lifetime and thus the whole assembly lifetime can be much improved by using hourglass shaped solder joints.

Having discussed the effects of solder bump height and shape on reliability, it becomes clear that the optimum solder bump design would be high standoff hourglass-shaped structure. The solder joints made by conventional solder bump technologies are in barrel shape which concentrate stresses and strains at the chip pad and substrate metallization interfaces and require large pitch to obtain reliable joint heights. In an effort to optimize solder bump connection geometry and shape, we have developed a novel stacked solder bumping technology for enhanced reliability of the joints. Figure 4 illustrates the detailed structures of conventional solder bumps as well as the triple-stacked solder bumps. As can be seen from the figures, a structure of triple-stacked solder bumps consists of an inner bump, middle bump and external bump. The stacked solder bumping technology controls the solder bump geometry and shape easily and offers the improvement in solder joint height and shape. Figure 5 shows some solder bumps fabricated by the triple-stacked solder bumping process. We can see that this solder bumping process provides the opportunity of achieving different solder bump heights and shapes for different chip pad size and solder joint volume design.
This new "stacked solder bumping" technology is an extension of the well-established solder bump technology. This assembly technology offers several advantages. It offers an improvement and choice in the solder joint height and shape. It can not only increase the standoff height, but also achieve different shapes, including hourglass shape, by keeping the other design parameters such as pad size unchanged. The process is compatible with the surface mount technology and it is feasible for volume production. It is potentially low cost. In the conventional solder bump technology, because of the tolerances on the height of the solder bumps and size of pads, and the warpage of the package and substrate, it is a potential problem that not all the bumps may be in contact with their pad mate. However, this problem is accommodated in this stacked solder bumping technology as described in the following sections. The major disadvantage of this stacked solder bumping process is that it involves additional steps of printing and reflow.

The melting point of the solder, the bumping process capability, manufacturability, and reliability are some of the major factors which influence the choice of solder compositions. Temperature hierarchy must be preserved at all times during the assembly processes to ensure structural integrity of the solder bump interconnect. Solder composition also has a significant influence on solder bump reliability. Eutectic Sn96.5/Ag3.5 solder has excellent characteristics and has shown great reliability improvement over eutectic lead-tin solder [2, 31]. Based on these principles, we selected the solder compositions of each layer. The inner bump is of Sn96.5/Ag3.5 alloy with a melting temperature of 221°C. The middle bump is of Sn10/Pb90 solder with a melting temperature of 268°C. The external bump is eutectic solder (Sn63/Pb37) with a melting temperature of 183°C. The height of this triple-stacked solder bump is in the range of 20-50 mil depending on different design and application.

Figure 4. (a) Single solder bump fabricated by the conventional way; (b) Triple-stacked solder bump fabricated by the new technique.

Figure 5. Microphotograph of solder joints that have different heights and shapes fabricated by triple-stacked solder bumping technology.

3. Stacked Solder Bumping Process

The stacked solder bumping process consisted of three basic processes: Stencil printing, solder ball placement and reflow. Figure 6 shows the stacked solder bumping process. Stencil-printing process involved three steps. The solder paste was first pushed into the holes in the stencil by a squeegee and it made contact with the bond pads on the chip. Then, the paste was transferred to the chip pads while the stencil was retracted. Finally, the paste bump was prebaked in order to retain its shape during the next process. The solder ball placement process was quite straightforward. First a stencil was placed on top of the chip, and then commercial solder balls were dropped in the windows of the stencil and stick to the prebaked inner solder bump. In our research, the power chip pad size was 1.1
mm. According to our design of forming hourglass-shaped solder joint, 35 mil (0.9 mm) diameter solder balls were used. The last process of the stacked solder bumping was reflow. As we stated earlier, the inner solder bump was Sn96.5/Ag3.5 alloy with a melting temperature of 221°C. The solder ball was of Sn10/Pb90 solder with a melting temperature of 268°C. During the reflow process, the stacked solder bump was heated to 250°C. This temperature is above the inner solder bump melting temperature, but below the solder ball melting point. When the inner solder melted, it formed metallurgical bonds with both the chip bond pad metallization and the top solder ball. In order to reduce oxidation of the surfaces and minimize voids contents in the interfaces, the reflow operation was performed in a reduced atmosphere of nitrogen-hydrogen. Figure 7 (a) shows the stacked solder bumps on IGBT pads and Figure 7 (b) is the magnified photograph of a solder bump. The under bump metallization (UBM) is Ti/Ni/Ag.

4. Flip Chip Assembling Process

This process involved flip chip bonding and flip chip underfill. First a photoimagable solder mask was applied to the prepatterned substrate with conventional screen-printing. Photolithography allowed definition of openings in the solder mask around all the chip site pads and surface mount footprints on the substrate. This also offered the alignment mark for the chips. Figure 8 shows the triple-stacked solder bump bonding process. During flip chip bonding, external solder paste was first stencil-printed on the substrate. Then the bumped die was aligned and attached to the printed solder paste on the substrate. Lastly the assembly was heated so that the bottom solder melted and formed a metallurgical bond with the bond pad. Also as the melting of the bottom solder occured, the surface tension for the melted solder and gravity of the bumped die caused the bumped die to be pulled down, thus allowing all the solder bumps be connected. Again, the temperature hierarchy principle must be obeyed. The external bump was eutectic solder (Sn63/Pb37) with a melting temperature of 183°C. We reflowed the assembly at 210°C. As thus, we formed triple-stacked hourglass shape solder bump interconnect. Figure 9 shows the flip chip assembly before underfill. Figure 10 is the cross-section image of the triple-stacked solder bump.
Figure 8. Triple-stacked solder bump bonding process.

Figure 9. Flip chip on substrate assembly before underfill.

Figure 10. SEM cross-section view of a triple stacked solder bump.

After flip chip bonding and cleaning, electrical test was performed prior to underfill of the solder joints. During underfill, the assembly was placed on a hot plate at 90 °C. Underfill was placed around two chip sides in an “L” pattern. After some time, the gap was completely filled and then the underfill was cured. At this stage, the assembly was quite robust and ready for the next level packaging.

5. Reliability Assessment and Failure Analysis

The interfaces of the triple-stacked solder joints were examined using scanning electron microscopy energy dispersive X-ray analysis (SEM-EDX) for the integrity of the joint. The adhesion strength of the solder bump to bond pad was characterized and analyzed. The reliability of the triple-stacked solder bump as well as the conventional solder bump was evaluated using thermal cycling test and the integrity of the joints were monitored and investigated by the change of forward voltage (our test samples were IGBTs and diodes) and microphotographs.

Interfaces Characterization: Generally, the weakest interface in flip chip assembly is the interface between solder bump and silicon chip. The triple-stacked hourglass solder bump was developed to reduce the stress concentration at the interface. However, failure could come first from the triple-stack solder bump since it is composed of three bumps and has two interfaces. It is important to examine the interfaces. Figure 11 (a) is the SEM picture of the triple-stacked solder bump and Figure 11 (b) is the corresponding EDX mapping. Figure 12 shows the high-magnified interface (the box in Figure 11 (a)) between middle solder bump (Sn10/Pb90) and external solder bump (Sn63/Pb37). The left EDX pictures are the Pb element EDX mappings, while the right ones are the Sn element mappings. SEM results reveal that the interfaces between the three solder bumps for as processed triple-stacked...
solder bump structure are consistent. EDX results show that the boundaries of different solder compositions are obvious and it indicates that there is no much diffusion between the different solder compositions and form alloy.

![Figure 11. (a) SEM picture; (b) EDX mapping of the triple-stacked solder bump structure.](image)

![Figure 12. (a) SEM picture; (b) EDX mapping of the interface between middle solder bump (Sn10/Pb90) and external solder bump (Sn63/Pb37).](image)

**Adhesion Test:** The adhesion between the interfaces of the Al pad and UBM film, and the UBM film and solder bump, is critical to the D²BGA chip-scale package and the 3-D power module since these are the most vulnerable interfaces. It is very important to test the bonding strength of the solder bumps. The testing was performed on an Instron machine controlled through its GPIB interface using LabVIEW software. The samples were loaded at a crosshead displacement rate of 1 mm/min. Utilizing LabVIEW software, the load at break was recorded. The load value is supposed to increase with the increasing displacement linearly within the elastic region. As the load versus displacement curve deviates from linearity, it indicates that cracking begins to occur in the joints. When the load reaches a maximum and begins to decrease rapidly, the solder joint has achieved a maximum loading level and completely breaks. Tensile tests were conducted on both stacked solder bump and conventional single solder bump for comparison. Also both in-house sputtered UBM Cr/Cu and vendor supplied solderable devices (Ti/Ni/Ag UBM) were used. In order to do the testing, a thin wire was soldered on top of the solder bumps. Figure 13 shows the test samples.

![Figure 13. Adhesion test samples. (a) Stacked solder bump; (b) conventional single solder bump;](image)

Our test results showed that the failure mode for all the samples was interfacial fracture. However, the adhesion strength for different samples was quite different. Figure 14 and Figure 15 are typical load vs. displacement curves for solder bump on IGBT pads. We can see that the adhesion strength for solder bumps on in-house made UBM device pad are much lower than that on vendor supplied solderable device pad. Interestingly, the adhesion strength for conventional single solder bump and stacked solder bump is almost the same for in-house
made solderable device. However, the adhesion strength for conventional single solder bump and stacked solder bump is quite different for vendor supplied solderable device. This could be understood after we investigated the fracture interface. From optical microscope observations, we found that the fracture occurred at the UBM/Al pad interface for in-house made solder devices, while for vendor supplied solderable devices, the fracture happened at the interface between UBM and solder bump (or probably between the UBM layers). This indicated that the quality of the in-house made UBM was not as good as that of the vendor supplied. From the curves in Figure 15, we can see that the adhesion strength for stacked solder bump is about 35 Newton, while that of single solder bump is about 25 Newton for vendor supplied solderable devices. We believe the reason for this difference is that the stacked solder bump has lower stress concentration at the corners of the UBM/solder interface than single solder bump. As we discussed before, there are stress singularities at the contact corners of the solder bumps. The stacked solder bump has smaller contact angle and thus less severe singularity, as illustrated in Figure 16. On the other hand, the adhesion strength difference verified that the stacked solder bump structure reduced the stress concentration at the corners of the solder bump.

![Figure 14](image1.png)

**Figure 14.** Typical load-displacement curve under tensile test for (a) Conventional single solder bump; (b) stacked solder bump on in-house sputtered Cr/Cu UBM device pad.

![Figure 15](image2.png)

**Figure 15.** Typical load-displacement curve under tensile test for (a) Conventional single solder bump; (b) stacked solder bump on vendor supplied solderable device pad (Ti/Ni/Ag UBM).

![Figure 16](image3.png)

**Figure 16.** Pictorial representation of stress concentration for (a) Conventional single solder bump; (b) stacked solder bump.
Thermal Cycling Test: Thermal cycling test is one of the most important tests used to assess the reliability of solder bump interconnection. In order to investigate the effects of solder bump geometry and shape on reliability, packaged power modules with both triple stacked hourglass-shaped solder bump and conventional barrel-shaped solder bump as chip-level interconnection were conducted thermal cycling test. For the same interconnection structure, the modules with underfilled and nonunderfilled chips were tested. The thermal cycling temperature range was 0°C and 100°C with the rate of 2 cycles/hour. The changes of collector emitter voltage drop $V_{CE(sat)}$ of IGBT devices and/or forward voltage $V_F$ of diodes were used as the evaluation criterion. SONY Tektronix 371 programmable high power curve tracer was used to measure the $V_{CE(sat)}$ and $V_F$. The modules were tested (output and breakdown) for functionality dynamically during the power cycling at both high temperature (100°C) and low temperature (0°C). After every 200 thermal cycles, the modules were systematically tested. When there is any crack in the solder bump, the voltage drop $V_{CE(sat)}$ and forward voltage $V_F$ will increase.

Figure 17 shows the typical forward voltage changes for the four different configurations of packaged power module as a function of the number of cycles. The values of the forward voltage for different power modules were normalized in order to compare. All of the test samples were functional and no evident failure could be detected until 3000 temperature cycles. However, after 3400 thermal cycles, there was obvious increase in the forward voltage value for the samples with barrel-shaped nonunderfilled solder bumps as the interconnection. After 5400 cycles, the barrel-shaped nonunderfilled solder bumps broke completely. Figure 18 (a) shows a fractured solder bump. As we can see, the failure occurred at the interface between the solder bump and the silicon chip. The forward voltage value of the samples with hourglass-shaped nonunderfilled solder bumps had noticeable increase at about 6000 cycles and it gradually increased afterwards. By 8600 cycles, it increased almost 20%. We examined the solder bumps using optical microscope every 200 thermal cycles after we noticed the $V_F$ increase and at 6400 cycles we observed obvious crack initiation in some of the solder bumps which were located at the chip corners, as shown in Figure 18 (b). Also we can see that the crack location of the hourglass-shaped solder bump was inside the stacked solder bump (not the interface of the chip and solder bump), which was different from that of the barrel-shaped solder bump. Furthermore, underfill improved the reliability of both barrel and hourglass solder bumps. The experimental work is still going on and more samples are under evaluation.

![Figure 17. Typical forward voltage changes of power modules versus number of thermal cycles](image)

![Figure 18. Microphotographs of (a) barrel-shaped solder bumps failed completely after 5400 thermal cycles; (b) hourglass-shaped solder bumps had crack initiation after 6400 thermal cycles.](image)
6. Conclusion

A new stacked solder bumping technology has been developed to fabricate hourglass-shaped high standoff solder bumps. This technology is a feasible low cost solution to achieve high solder bump reliability. The preliminary experimental results revealed that the triple-stacked hourglass-shaped solder bump produced by the stacked solder bumping technology was much more reliable than the traditional solder bump.

A systematic reliability evaluation and failure analysis of the stacked solder bump as well as the conventional single solder bump is being conducted and will be reported.

Acknowledgments

We thank Dr. Chendong Bai, Sihua Wen, Dan Huff and Shu Guo for their assistance and useful discussion. This work was supported in part by the ERC Program of the National Science Foundation under Award Number EEC-9731677.

References


