State of The Art in Semiconductor Package for Mobile Devices

Jin Young Kim · Seung Jae Lee

Amkor Technology Korea

Abstract

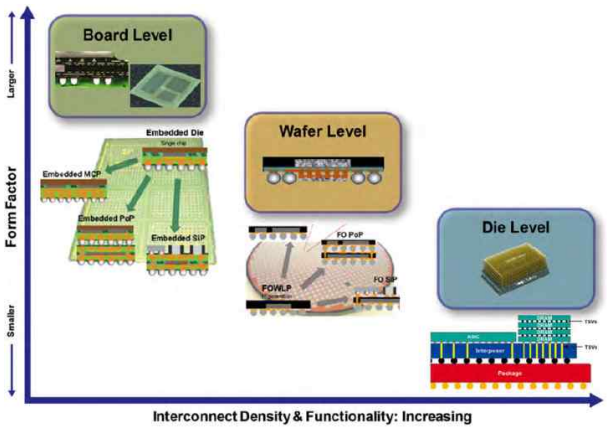
Over the past several decades in the microelectronics industry, devices have gotten smaller, thinner, and lighter, without any accompanying degradation in quality, performance, and reliability. One permanent and deniable trend in packaging as well as wafer fabrication industry is system integration. The proliferating options for system integration, recently, are driving change across the overall semiconductor industry, requiring more investment in developing, ramping and supporting new die-, wafer- and board-level solution. The trend toward 3D system integration and miniaturization in a small form factor has accelerated even more with the introduction of smartphones and tablets.

In this paper, the key issues and state of the art for system integration in the packaging process are introduced, especially, focusing on ease transition to next generation packaging technologies like through silicon via (TSV), 3D wafer-level fan-out (WLFO), and chip-on-chip interconnection. In addition, effective solutions like fine pitch copper pillar and MEMS packaing of both advanced and legacy products are described with several examples.

Ⅰ. Introduction

System integration is the key issue driving the packaging sector, with integration at the board, wafer, and die level happening simultaneously. Most of major OSATs put all resource and capital investments into packaging technology for the smart phone now to build up a rigid infrastructure for the future. Although disruptive changes are afoot all these areas, the basic platforms and accumulated learning for handling common issues like thinning, warpage, and chip/package inter-connection are already in place.

To track the trend of integration and miniaturization and highly functional electronic devices, multiple efforts are being droved by OSATs as shown [Fig. 1]. At the die level, through-silicon via will at last become the big play for advanced system integration in two or four years, with silicon interposers for integration of application processor (AP) and advanced memory, wide I/O, for future smart phone and tablet PC. Technology that is just coming next will likely be transitional technologies for die-to-die stacking on a single platform. One solution now in development for possible production near term is a hybrid die stacking technology that ex-



[Fig. 1] Board, wafer, and die level trends toward system integration. (Courtesy of Amkor Technology)

tends flip chip towards the TSV space to increase band- width and decrease signal latency. It is called chip-on- chip, face-to-face technology.

At the board level, integrated passives in substrate are now mainstream while embedded active die are less common as PCB technology is still being improved. Even though embedded die packaging allows higher density and more I/Os, the less tightly controlled PCB technology means higher yield loss. While efforts to get higher yield and lower cost for embedded active die are being progressive, the primary approach for reducing this loss is to add the die at the very end of process in a cavity on the top side of the substrate.[2],[3] However, unfortunately, making cavities in laminate substrate remains expensive. To address the needs of next generation board integration, substrate suppliers will also have to step up their investment in development to extend the technology from 15um down to possibly < 5 um lines and spaces. This should require improvements in lithography process as well as materials and need a lot of investments. As a result, integration on silicon interposer or wafer level packaging technology is more viable option for many applications.

At the wafer level, wafer-level fan-out (WLFO) packaging is on niche market for specific devices like RF and high frequency application and 3D extension as platform solution like POSSUM^{TM}, low level interposer, and wafer level package-on-package (PoP).

In this paper, cutting edge trends that are advanced technology like through silicon via and coming next technology of chip-on-chip process, platform technology of wafer-level fan-out (WLFO) in packaging industry for mobile devices are introduced. In addition, effective technology to support advanced products like fine pitch copper pillar and MEMS devices are introduced.

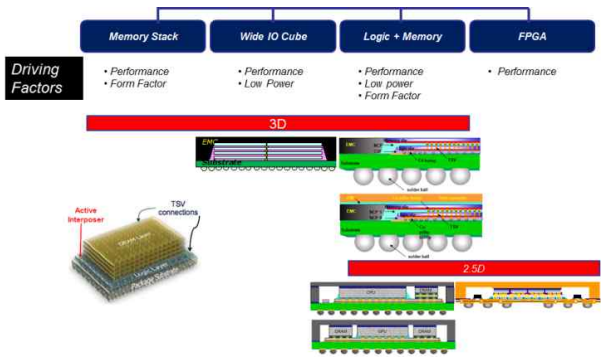
Ⅱ. Through Silicon Via (TSV)

Through silicon via (TSV) provides the advantages of high density, high bandwidth connections between side-by-side die. It has also served well to accelerate the migration of packaging technology from single die platform to multiple die options. The 3D TSV products are long term goal of many products and applications to maximize package level integration. Major application of TSV is described on [Fig. 2].

Driving force of TSV developments could be distinguished 4 sectors that are high performance memory stack for network server, wide I/O memory cube, logic & memory integration, and FPGA for large body flip chip.

Specifically, in mobile application, 3D stacking and 2.5D multiple die placements, as shown [Fig. 2], are primarily represented by mobile application spaces for high end products. The purpose of use of TSV in mobile application is integration of logic, application processor (AP), and memory.

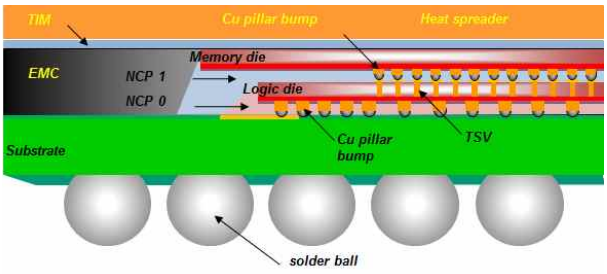
In addition, functionally, there are pros and cons in



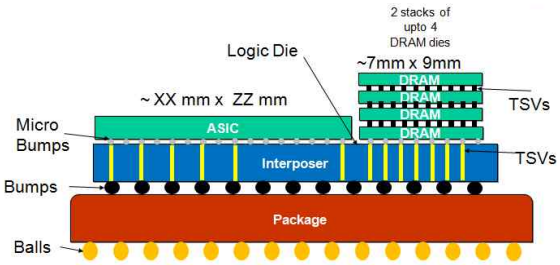
[Fig. 2] Major application of TSV. (Courtesy of Amkor Technology)

both cases. In case of 3D stacking, it would have lower cost than 2.5D but it should be expected degradation of thermal performance. While 2.5D approach would have better thermal performance than 3D, it may have higher cost than 3D, process issues on interposer interconnection to substrate due to warpage, and more process steps like passive components attach.

To achieve this technology, fine pitch copper pillar



(a) 3D stacking TSV



(b) 2.5D TSV

[Fig. 3] Example of 3D and 2.5D TSV. (Courtesy of Amkor Technology)

and microbump product developments provides the foundations for TSV assembly development.

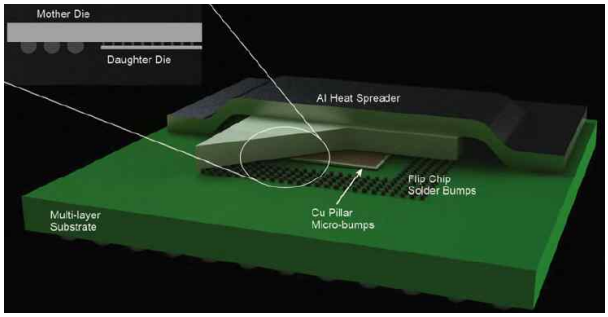
One more thing that need to be prepared for TSV process in OSAT is MEOL wafer processing. Since the OSATs and foundries have learned that transferring thin wafer caused lots of issues, wafer bonding and debonding need to be done at the same site. That leaves the OSATs with the challenge of investing significantly and learning to master the complex CMP fab process for TSV via reveal.

According to review TSV packaging process, TSV implementation requires considerable commitment in terms of expensive capital investment in equipment and tooling from wafer handling to stacked assembly and test. Fabrication steps, such as etching through the holes on the back of the wafer, also add significant cost to the final assembly. However, it should be evident that TSV ultimately provide the coveted path towards miniaturization, increased functionality, and more storage capacity at the package level for future mobile devices.

Ⅲ. Chip-on-Chip Technology

Most people may agree that TSV is most active areas of research and development in semiconductor packaging over the next 5～10 years in anticipation of its commercialization. Before becoming TSV era, OSATs would prepare just next coming technologies. Chip-on-chip technology as shown in [Fig. 4], is one of competitive in their target market today.[1] In order to get the closest connection between the active circuitry on each die without the expensive TSV processing, the face-to-face configuration is proper approach.

The stacked die configuration describes two or more devices assembled face-to-face where a smaller die is

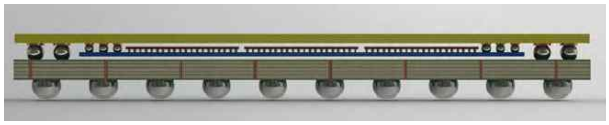


[Fig. 4] Conceptual illustration of a POSSUM^{TM} assembly. (Courtesy of Amkor Technology)

nested within I/O-free areas of the larger die. The larger of the two dies is referred to as the mother die and the smaller one is called the daughter die. As shown in [Fig. 4], the daughter die is flip chip mounted onto the face of the mother die so as not to interfere with the mother's surrounding flip chip bump pattern.

The chip-on-chip, face-to-face design is very attractive to high performance as well as integration because it provides a close coupling electrical signaling path between mother and daughter die. From this structure designer can also achieve less parasitic elements like inductance and resistance that prevent high speed signaling from die to die than wire bonding. Only concern on chip-on-chip technology is finding out thermal dissipation path. However, it can be easily improved using integrated heat spreader (HIS).

[Fig. 5] illustrates a double chip-on-chip with multi die stack package. This design is actually defined by two levels of nesting die. The daughter die is flip chip attached to the mother die which is then attached to another silicon die instead of a substrate or board. This largest die is than flip chip attached to the substrate or board. The levels of nesting are dependent upon the overall profile of each assembled die, including warpage factors and the inclusion of any added materials such as



[Fig. 5] Double POSSUM^{TM} multi-stacked die configuration without the use of TSVs. (Courtesy of Amkor Technology)

not-conductive pastes, underfills, mold compounds, etc.

There is much interest from the microelectronics community because of the modular approach to die integration, the ability to preserve speed and bandwidth without introducing excessive latency or parasitics, and the lower cost involved to produce effective 3D solutions. And it is being sought in mobile markets as diverse as logic and memory, MEMS (timers, accelerometers, gyroscopes), optoelectronics, and microcontrollers. This is because it uses the existing chip attach or thermocompression (TC) infrastructure and can be ramped to HVM without complex wafer handling processes, making the cost less expensive than TSV approaches. Although it is not a TSV replacement, it enables high performance integration today and will be an adjunct to full TSV integration tomorrow.

Ⅳ. Wafer-Level Fan-Out (WLFO)

During recent several years, wafer-level fan-out (WLFO) is designed to provide increased I/O density within a reduced footprint and profile for single die one-layer RDL packaging applications at lower costs. This baseline structure is experiencing increased interest due to its ability to also integrate mid-range I/O density to support mobile data modems, power management ICs, CODECs, RF switches, and power amplifier for mobile device.

Two-Dimensional (2D) WLFO is well-established as a robust and reliable wafer-level package (WLP) technology for electronic devices and as a viable alternative to conventional laminate-based and wafer-based packages. [Fig. 6] illustrates a cross section of a typical 2D WLFO structure.

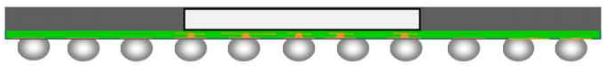
The fundamental WLFO technology is based on the embedding of die into a molded wafer (also referred to as 'wafer reconstitution'). The molded wafer is then processed through standard wafer-level packaging processes to create the final package structure. The active surface of the die is coplanar with the mold compound, allowing for the 'fan-out' of conductive copper traces and solder ball pads into the molded area using conventional redistribution layer (RDL) processing. [Fig. 7] shows a cross section photo of the die and mold transition area.

However, the 2D single die and dual die WLFO devices using wafer lithography process is relatively expensive, Accordingly, to present 2D WLFO to expand wide market portion, one of major OSATs are actively researching on development of laser direct write process.[1],[4],[6]

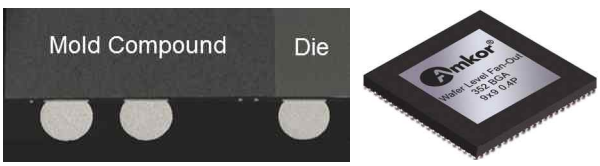
Positive development and qualification results indicate that WLFO has the ability to extend itself beyond simple 2D structures to meet the needs of more advanced applications and device technologies. More recently, WLFO has shown its ability to extend itself into more innovative and advanced structures to support the increased functionality and performance requirements of next generation mobile and networking devices. By applying enabling technologies, such as Thru-Mold Via (TMV®)[5] Fine Pitch Copper Pillar (FPCuP) bumping and Chip-on-Chip (CoC) bonding, WLFO extends its applications and benefits to the area of 3D Packaging. Two such structures are 3D PoP (Package-on-Package) WLFO and Face-to-Face (F2F) WLFO.

2D wafer level fan out (WLFO) technology can be possibly expanded to Z-direction connected 3D PoP (Package on Package) Structure as shown [Fig. 8] using advanced laser and via fill process.

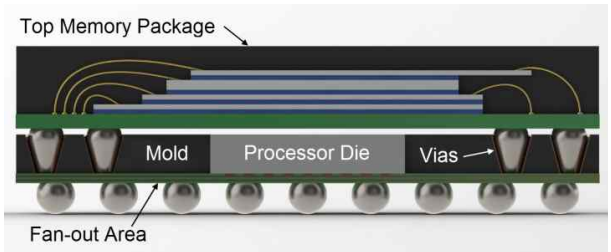
The front side (i.e., active die side) RDL process for 3D PoP WLFO is nearly the same with that of standard 2D WLFO. Moreover, a laser drilling (or equivalent) process is used to expose an RDL feature from the back side (i.e., the mold side). [Fig. 9] illustrates the key process steps for the 3D PoP WLFO fabrication process.



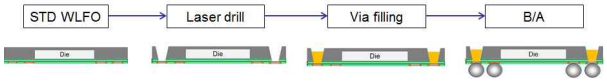
[Fig. 6] 2D WLFO structure.



[Fig. 7] SEM of WLFO cross section and 2D WLFO pacakge.



[Fig. 8] 3D WLFO PoP package structure.

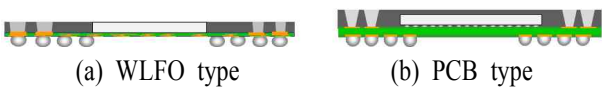


[Fig. 9] Key process flow steps for 3D PoP WLFO.

3D PoP WLFO has additional technical merits due to using wafer level RDL technology. First of all, RDL process only needs tens of micron thickness instead of several hundreds of thickness PCB wouldn't need to be used and thinner thickness would give better electrical performance due to less parasitic elements. In addition, the elimination of additional interconnection features, such as wire bonds or flip chip bumps, provides the opportunity for cost reduction. One more thing is that via pitch of 3D WLFO PoP is relatively smaller than that of PCB based product.[7] Consequently, extremely thin package could be coming true by wafer back-grinding after bumping process as shown in [Fig. 10].

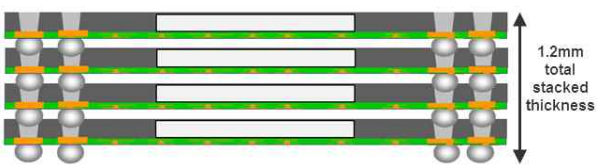
These kinds of advantages would be used on mobile device like multi-stack structure as shown on [Fig. 11]. By controlling individual package thickness under 300 um, as shown in <Table 1>, final thickness of 4-stack package would be achieved with 1.2 mm height.

Another attribute of 3D WLFO is that, through the use of through mold via (TMV®) technology, the active die side can be oriented face-up and exposed to the environment – an ideal condition for biometric sensing applications. Fine pitch vias are used to make connections between the front (die) side and back (mold) side, to translate the solder ball pads to the back side of the package. This structure is illustrated in [Fig. 12].



(a) WLFO type (b) PCB type

[Fig. 10] Comparison of 3D PoP structures.

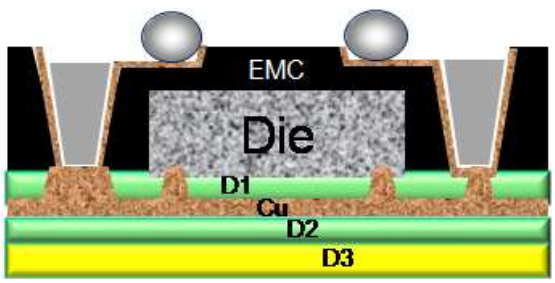


[Fig. 11] Example of 3D PoP WLFO memory stack.

<Table 1> Package dimension. (single PoP) (unit: mm)

Item	WLFO
Mold cap	0.15
Substrate (WLFO : 2 layer RDL)	0.05
BGA (0.4 mm pitch)	0.10
Total thickness (max)	0.30

The core technology of WLFO 3D PoP is the TMV® process that drills vias through epoxy mold compound (EMC) using laser ablation. The vias are then filled with a conductive material to establish a reliable electrical connection with the underlying RDL feature. One of the constraints of the TMV® structure is its limitation in creating increasing finer pitch and deeper via geometries. Process characterization is required to ensure the via structure can endure the thermal and mechanical stresses of the TMV® process. The intense localized heating during the laser drilling process, coupled with a high plasma gas environment, can cause oxidation and contamination on the targeted copper RDL pad on the front side of the package. Process optimization is critical to ensure a robust TMV® process and a contamination-free and oxide-free electrical interface. In addition, EMC materials are highly concentrated with silica filler. When sub-

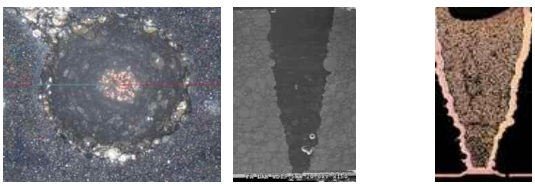


[Fig. 12] Illustration of 2.5D WLFO.

jected to laser ablation, the filler breaks into very fine particles that can contaminate the surface of the wafer. To minimize the risk of contamination, special cleaning methods are applied after the laser ablation process. [Fig. 13] illustrates the laser drilling process and shows examples of drilled and filled vias.

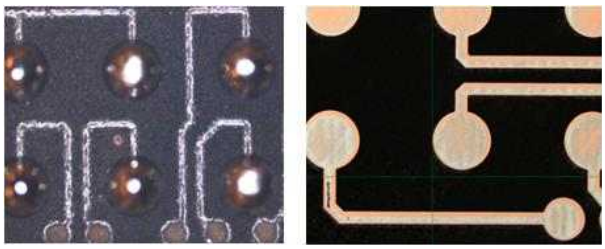
The top package in a TMV® PoP stack typically establishes its connection by nesting its solder balls directly on top of the underlying package's TMV® structure (as previously illustrated in Fig. 14). In cases where redistribution of the TMV® pattern is needed on the mold side of the package, direct laser patterning techniques can be selectively applied, either on a polymer layer or directly on the EMC surface. [Fig. 14] shows redistribution patterns connecting the TMV® structures to the BGA pads by using direct laser ablation on either the EMC surface or on a polymer layer.

Another innovative 3D structure is Face-to-Face WLFO. This fan-out package, shown in [Fig. 15], has a depopulated BGA pattern within which a second die is

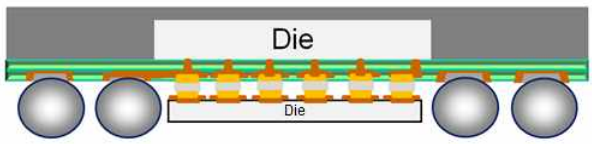


(a) Laser via drilling (b) After via filling

[Fig. 13] Key process flow steps for 3D PoP WLFO.



[Fig. 14] Redistribution layer and TMV structure on EMC (left) and on a polymer layer (right).



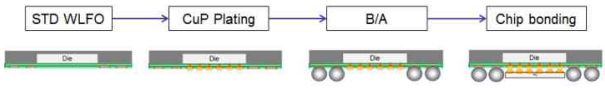
[Fig. 15] F2F WLFO structure.

connected directly to the underlying WLFO RDL structure. This direct chip attach methodology results in a small form factor package that provides very low latency signal path for high speed data transfer between memory and logic components.

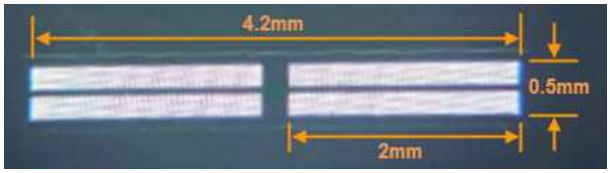
General process flow of F2F WLFO is briefly described in [Fig. 16].

The core technology (and key challenge) for F2F WLFO processing is pattern alignment. A molded wafer (as opposed to a standard silicon wafer) shrinks and expands due to the thermal stress it experiences during WLFO processing. In addition, the actual die stepping distance of a molded wafer has a more significant deviation from the design value due to the die attach machine's placement tolerance. Consequently, design, material, and process optimization is needed to provide the most robust die to die bonding results.

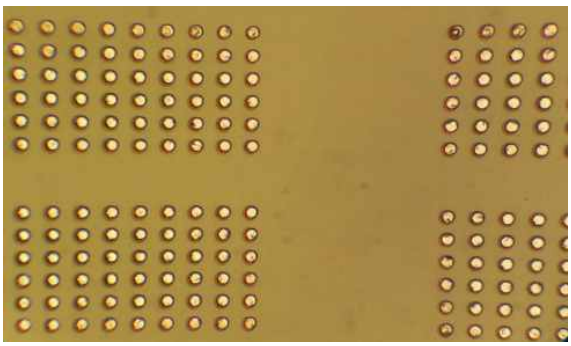
During the redistribution processing of a F2F WLFO structure, optimized conditions must be established that control the distortion and warpage of the molded wafer within allowable tolerances. This is a key element to ensure the successful alignment of the CoC interconnection during the chip mounting process. [Fig. 17] shows the actual F2F interconnect area on the molded die. [Fig. 18] is a picture of the molded die's fine pitch (40



[Fig. 16] Overall process of F2F WLFO.



[Fig. 17] F2F interconnect area on WLFO.



[Fig. 18] Fine pitch (40 um) copper pillar F2F interconnection.

〈Table 2〉 3D PoP and F2F WLFO reliability results.

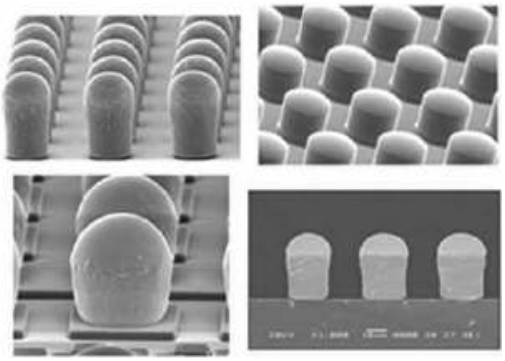
		3D PoP	3D F2F
MSL	L3 260℃	PASS	PASS
Temp cycle condition B	500 cycles 1,000 cycles	PASS PASS	PASS PASS
uHAST	85℃/85 % RH	PASS	PASS
High temp storage	150℃ 1,000 hours	PASS	PASS
Board level reliability	Temp cycle & drop shock	Ongoing	Ongoing

um) interconnect pattern. This F2F WLFO test vehicle's overall package size is 14×14 mm, the molded die and CoC die are both 8×8 mm, and the final F2F package thickness is 250 um. The number of interconnects on the CoC die is 1,200 with a pitch is 40 um.

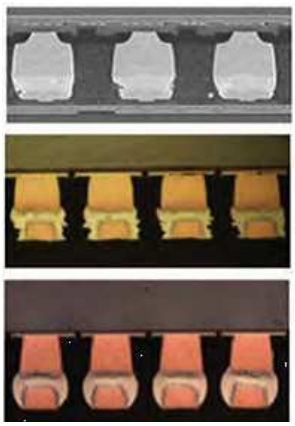
Reliability have been performed on 3D PoP and F2F WLFO package. The status and results of these tests are shown in 〈Table 2〉.

Ⅴ. Fine Pitch Copper Pillar

A big part of making a platform meet the needs of a range of application to drive it to high volumes is of course to drive down its costs. For example, fine pitch copper pillar have been developed for targeting low cost from the beginning to compete with wire bonding. That meant figuring out how to eliminate the passivation/repassivation layer and make the bumps directly on the Al pad, to reduce the process to a single mask step. However, the recent explosion of smartphones and tablets into the mobile market with high density I/O, the results of development is major reduction in bump pitch. Actual processed example is illustrated in [Fig. 19].



(a) Fine pitch copper pillar before bonding



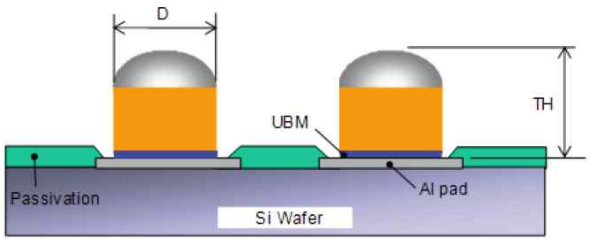
(b) Fine pitch copper pillar after bonding

[Fig. 19] Fine pitch (40 um) copper pillar F2F interconnection.

Bump pitches can be reduced to as low as 50 um in line and 40/80 um staggered. Benefits of such small bump pitches are that substrate layer count can be reduced and lowering overall package heights and material costs.

Copper pillar design rule and pad design guidelines are described in [Fig. 20, 21] and <Table 3, 4>.[1]

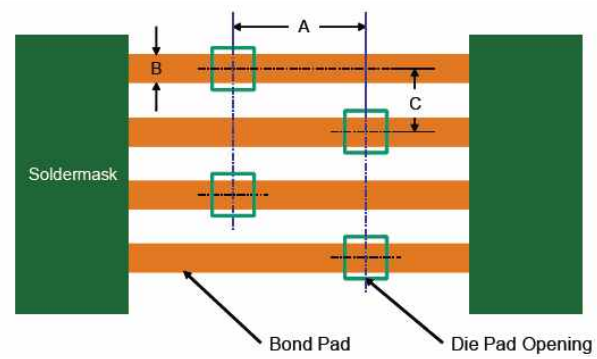
Copper pillar technology has many benefits, includeing excellent electromigration performance for high current-carrying applications and miniaturization. But, a major advantage of copper pillar bumping exists for chip desiner who are transtioning from wire-bond packaging to flip-chip packaging. Copper pillar assembly allows the chip design to maintain peripheral-based signal routing, without time-consuming redesign, thus accelerating time-to-market and reducing design costs.



[Fig. 20] Copper pillar design rule.

<Table 3> Range of copper pillar size.

Feature	Dimension
Cu pillar diameter (D)	20～50 um
Total height (TH)	30～45 um



[Fig. 21] Copper pillar pad design guideline.

<Table 4> Copper pillar pitch and pad size.

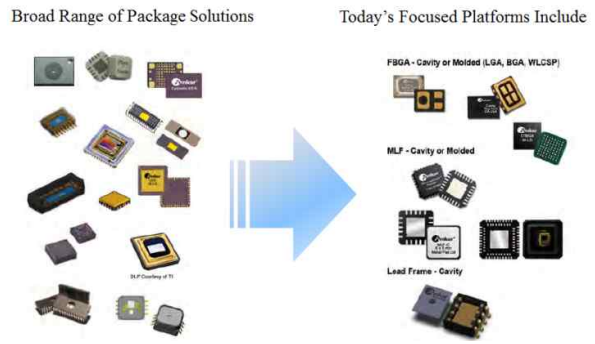
General design rules		Pitch				
		60	50	45/90	40/80	30/60
A	Row to row pitch			90	80	60
B	Bond pad width	30	25	22	20	
C	Trace pitch	60	50	45	40	30

VI. MEMS Packaging

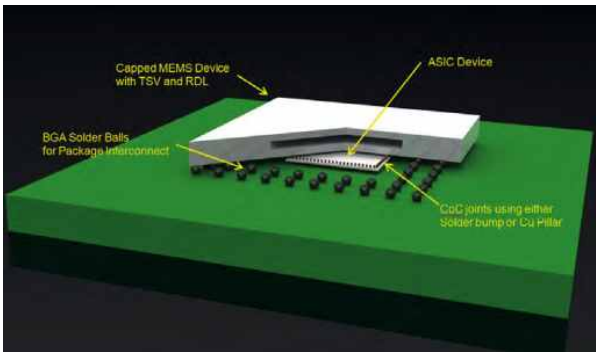
OSATs are addressing the market's needs for flexible standard package platforms for MEMS applications like accelerometers, gyroscopes, magnetometers and microphones that are used in mobile devices. Although the fabless MEMS makers have always been to adopt the standard packages offered by OSATs, the big IDMs who dominate the MEMS business with more custom designs are being driven by volume consumer markets to outsource more packaging, assembly and test. Fast growth and high packaging costs of custom MSMS packages make the market ripe for change and a ripe opportunity for those who can enable it. Even though IDMs have more defined requirements because they have years of expertise and more custom expectations, if they need high volume and fast cycle times with lower costs, it drives them to standard packages. In addition, outsourcing allows the IDMs to focus more of their capex on the front end, and leverage backend flexibility and risk mitigation with second sources for packaging. Transition from custom MEMS packaging to HVM standard is illustrated in [Fig. 22].

As shown in [Fig. 22], there are two kinds of standard package platforms[1], one is based on Micro-LeadFrame® (MLF), well known as name of QFN, the other is on a laminated or ceramic substrate, each topped with metal or a plastic lid that creates a cavity around the MEMS device for a low stress environment where the mold the molding compound is not in contact with the sensors. The cavity packages are good for multiple sensors that may have different stress levels.

New MEMS devices, in new kinds of packages, are also poised for growth. One of new approach for integrated MEMS devices is face-to-face chip-on-chip packaging, POSSUM™ configuration, as introduced in above chapter in this paper. The largest of the attached pair of chips is then flip-chipped onto a substrate or even directly to a motherboard as either a WLCSP or a BGA. Eliminating the need for a substrate to carry signals between die will significantly reduce the MEMS packaging cost. These approaches would be adopted for accelerometers, gyroscopes, and RF multi-chip assemblies as illustrated in [Fig. 23].



[Fig. 22] Copper pillar pad design guideline. (Courtesy of Amkor Technology)



[Fig. 23] Example of the face-to-face bonding of MEMS and ASIC devices creates a wafer level package. (Courtesy of Amkor Technology)

Though wire bonding still continues to dominate the MEMS packaging market, it now needs more discussion of flip-chip to decrease the size of the device and improve parasitic resistance and latency. Chip-on-chip configuration is a potential enabling technology toward low cost MEMS system integration.

Ⅶ. Conclusion

In this paper, state of the art in semiconductor package industry for mobile devices is briefly reviewed focusing on system integration and miniaturization solution. Through silicon via (TSV) is main stream for 3D system integration for next 5~10 years. Two different approaches of TSV, one is 3D-stacking, the other is 2.5D interposer application, are introduced. Chip-on-Chip technology, POSSUM™ is also reviewed to support cutting edge assembly technology for alternative forms to support high speed, high signal integrity, die-to-die communication. In addition, wafer-level fan-out (WLFO) with conjunction of 3D technology is also high lighted on mobile market now for specific applications like biometric sensors and 3D PoP structures that will be used in mobile products. In addition, technology of copper pillar is being more advanced to support current and future packaging application. Furthermore, recent packaging trend for MEMS devices using standard pack-

aging platform and advanced approach with POSSUM™ are described.

References

[1] http://www.amkor.com

[2] Nitesh Kumbhat, Fuhan Liu, Venky Sundaram, Georg Meyer-Berg, and Rao Tummala, "Low cost, chip-last embedded ICs in thin organic cores," *IEEE Electronic Components and Technology Conference 2011*.

[3] V. Sundaram, et al, "Chip-last embedded actives and passives in thin organic package for 1～110 GHz multi-band application," *IEEE Electronic Components and Technology Conference 2011*.

[4] Seung Jae Lee, V. et al, "Electrical characterization of wafer level fan out (WLFO) using film substrate for low cost millimeter wave application," *IEEE Electronic Components and Technology Conference 2010*.

[5] Jin Seong Kim, V. et al, "Application of through mold via (TMV) as PoP base package," *IEEE Electronic Components and Technology Conference 2009*.

[6] Seung Jae Lee, V. et al, "Electrical evaluation of wafer level fan out (WLFO) package using organic substrate for microwave application," *IEEE Electronic Systemintegration Technology Conference 2010*.

[7] T. V. Braun, et al, "Through mold vias for stacking mold embedded packages," *IEEE Electronic Components and Technology Conference 2011*.

≡ 필자소개 ≡

Jin Young Kim

received the B.S. degree in mechanical engineering from Hanyang University in 1990, Korea, and MS./Ph.D. degree in mechanical engineering from KAIST, Korea Advanced Institute of Science & Technology, Daejeon, Korea, in 1992, and 2002, respectively. He had been working for Amkor Technology Korea as leader of characterization group from 2000 to 2006 and move to Amkor ATK4 factory as team manager of RND branch from 2006 to 2008. His current research and development activities are focusing on cost effective 3D platform technology as leader of process development team.

Seung Jae Lee

received the B.S. degree in Electrical Engineering from Yeungnam University, Korea, in 2001 and the M.S. degree in Electronic Engineering from KAIST, Korea Advanced Institute of Science & Technology, Daejeon, Korea, in 2003, respectively. He has been working for Amkor Technology Korea, Seoul, Korea, as electrical and thermal characteristic engineer from 2003 to 2011. As a characterization engineer, his research interests include signal integrity, power integrity, and RF & millimeter wave design. From 2009, he is involved for development of wafer-level fan-out (WLFO) package as characterization engineer. His current research area is process optimization of excimer laser direct patterning, mechanical copper etching, laser via drilling and filling for 3D WLFO.