

HIGH PRODUCTIVITY UBM/RDL DEPOSITION BY PVD FOR FOWLP APPLICATIONS

Chris Jones, David Butler, Steve Burgess, Tony Wilby, Paul Densley
SPTS Technologies Ltd
Newport, United Kingdom
chris.jones@orbotech.com

ABSTRACT

Fan-Out Wafer Level Packaging (FOWLP) technology is an increasingly popular solution for obtaining high levels of device integration with a greater number of I/O contacts, at a lower cost.

In all FOWLP schemes, such as InFO (TSMC), SWIFT/SLIM (Amkor), eWLB (STATS ChipPAC, ASE, NANIAM) and many others, singulated die are embedded in epoxy mold compound (EMC). While EMC is a cost-effective material, it readily absorbs moisture, as do the polymer dielectrics used to electrically isolate adjacent metal interconnect lines. These materials must be degassed prior to physical vapour deposition (PVD) of under-bump metals and redistribution layers (RDL), otherwise metal to metal interfaces can be contaminated, impacting electrical performance of the device. With the low thermal budget of the EMC (<150°C), an effective degas requires low temperature and long process times that can significantly reduce throughput. SPTS has implemented a “*Multi-Wafer Degas*” solution to eliminate the “degas bottleneck” and ensure lowest contact resistance (Rc). This solution typically doubles the throughput compared with competing PVD systems, and is being used in high volume 300mm production.

Prior to UBM/RDL metallization, native oxide needs to be etched from the exposed metal contacts. In the SPTS Sigma[®] fxP PVD system, ICP technology is used to reduce the energy of ions at the wafer, and so limit the peak wafer temperature. In high density packages using multiple RDL, the upper interconnect level can present large exposed areas of metal to the pre-clean module, which can lead to a breakdown in RF coupling in ICP technology during etch. In addition, the presence of organic passivation such as PI or PBO can lead to premature particle failure due to the poor adhesion of organics to chamber furniture. This paper introduces a new pre-clean etch module, designed to provide ICP cleaning capability whilst overcoming the process and particle stability challenges posed by increased I/O contacts and organics. Data from systems in production will be presented, demonstrating the efficacy of the technology. The paper will show long time between chamber cleans in an ICP pre-clean module, and low & stable Rc data.

System modifications for coping with epoxy mold wafer warpage and cost reduction efforts are also discussed.

Key words: Fan-Out WLP, FOWLP, UBM/RDL, Epoxy Mold Compound, Rc.

INTRODUCTION

Our ability to create ever-smaller electronic devices that maintain or surpass the performance of their physically larger predecessors – exemplified by today’s wearables, smartphones and tablets – is dictated by many factors that extend well beyond Moore’s Law, from the underlying embedded components to the ways in which they are packaged together. With regard to the latter, fan-out wafer level packaging (FOWLP) is quickly emerging as the new die and wafer level packaging technique of choice, and is widely anticipated to underpin the next generation of compact, high performance electronic devices.

Whereas with conventional flip-chip WLP schemes the I/O terminals are spread over the chip surface area, limiting the number of I/O connections, FOWLP embeds individual die in an epoxy mold compound (EMC) with space allocated between each die for additional I/O connection points, avoiding the use of more expensive silicon real estate to accommodate a higher I/O count. Redistribution layers (RDLs) are formed using physical vapor deposition (PVD) and subsequent electroplating and patterning to re-route I/O connections on the die to the mold compound regions on the periphery.

Leveraging FOWLP, semiconductor devices with thousands of I/O points can be seamlessly connected via finely-spaced lines as thin as two to five microns, maximizing interconnect density while enabling high bandwidth data transfer. Significant height and cost savings are achieved via the elimination of the substrate.

With FOWLP today we have the ability to embed heterogeneous devices including baseband processors, RF transceivers and power management ICs in mold wafers, thereby enabling the latest generation of ultra-thin wearables and mobile wireless devices. With continued line and space reductions, FOWLP has the potential to accommodate higher performing devices including memory and application processors, positioning FOWLP to extend into new markets including automotive and medical applications and beyond.

Leading vendors implementing FOWLP today include Amkor, ASE, Freescale, NANIAM, STATS ChipPAC, and

TSMC, with TSMC being the most high-profile vendor given its widely-reported contract win to produce A10 processors for Apple's iPhone 7 – a deal said to be attributable in part to TSMC's mature FOWLP-based InFO technology.

According to a report entitled "FOWLP Forecast update 09/2015" published by research firm Yole Développement in September 2015 [1], the launch of TSMC's InFO format is expected to increase industry packaging revenues for FOWLP from \$240M in 2015 to \$2.4B in 2020. With a projected 54% CAGR, Yole expects FOWLP to be the fastest growing advanced packaging technology in the semiconductor industry.

PROCESSING CHALLENGES

In the FOWLP process flow, PVD is used in the RDL formation stage, depositing adhesion and seed metals in readiness for Cu electroplate formation of RDL lines and vias.

A typical PVD system for FOWLP comprises Degas, Etch and PVD module hardware (refer to Figure 1). Degas and Etch modules are used to remove contaminants and clean electrical contacts, and the PVD modules are used to deposit adhesion and Cu seed, ideally in dedicated chambers to prevent cross-contamination. At first sight the requirements for PVD may look no different to any other UBM/RDL applications but on closer inspection there are several aspects to the process flow that present fresh challenges to the user. It cannot be assumed that mainstream UBM/RDL sputter equipment can be used for FOWLP, at least not in a productive manner with high yields.

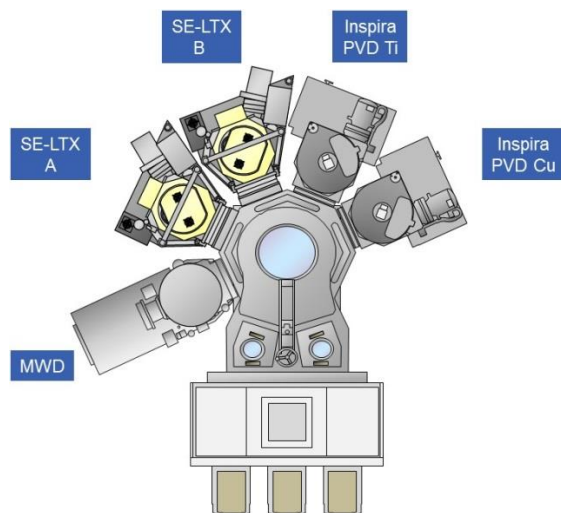


Figure 1. Sigma fxP PVD System for FOWLP configured with Degas, 2 x Pre-clean, 2 x PVD.

Low Heat, High Speed Processing

All fan-out wafers feature singulated die embedded in the EMC, with spin-on dielectrics surrounding the RDL. These materials present some unique challenges, including moisture absorption, excessive outgassing and a limited

tolerance to elevated temperatures. If not dealt with properly, contamination at the metal deposition stage can compromise electrical contact resistance and adhesion. Figure 2 shows typical RGA outgassing characteristics of a mold wafer. H₂O and CO₂ dominate.

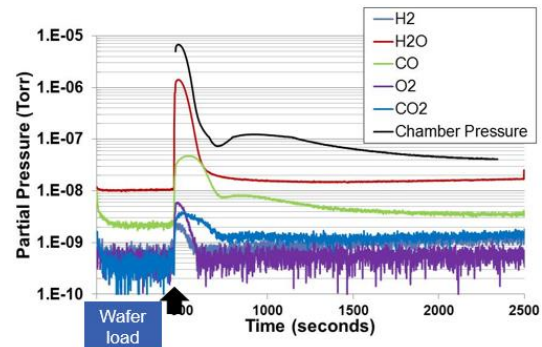


Figure 2. RGA Mass Scan of Mold Wafer.

Whereas conventional circuits built on silicon can withstand heat up > 400°C and can be degassed rapidly without impacting system throughput, the EMC and dielectrics used in FOWLP have a heat tolerance closer to 120°C. Temperatures exceeding this low threshold can cause decomposition and excessive wafer warping. Degassing wafers at such low temperatures naturally takes a longer amount of time, and can drastically reduce the throughput of a conventional sputter system. The data in Figure 2 shows that for a mold wafer it can take up to 30 mins for contamination levels to return to pre-load values.

Multi-wafer degas (MWD) technology has emerged as a compelling solution to this problem, enabling up to 75 wafers to be degassed at 120°C in parallel before being individually transferred to subsequent process steps, without breaking vacuum. With this approach, wafers are dynamically pumped under clean, high vacuum conditions, with radiation heat transfer warming wafers directly to temperatures within the operating budget for packaging applications. The use of high vacuum radiative heating also prevents cross-talk between wafers from a contamination perspective.

Each wafer can spend up to 30 minutes inside the MWD, but because they are processed in parallel, a "dry" wafer is outputted for metal deposition every 60 to 90 seconds, at a rate of between 30 to 50 wafers per hour. This approach increases PVD system throughput by 2-3 times compared to a single wafer degas processing technology, and as materials emerge with even lower thermal budgets based on increased passivation thickness, longer degas times can be accommodated with no impact on throughput.

Importance of Degas Under Vacuum

If a mold wafer isn't degassed sufficiently prior to pre-clean it produces high levels of outgassing that can affect plasma stability during etch, and film quality (Rc) during subsequent sputter deposition. The vacuum conditions in

which the degas takes place plays an important role. A test was performed measuring outgassing rates of plain silicon and silicon with organic material present. Wafers were subject to three test conditions: rate of rise test in a vacuum chamber after no degas, after degas in-situ (kept under vacuum), and after degas ex-situ (vented to atmosphere). Results in Figure 3 show that substrates with organic material degassed ex-situ showed outgassing characteristics similar to wafers that received no degas. Absorption of moisture under atmospheric conditions is rapid. This result shows that the best location of a PVD degas station is within the vacuum back end of the system.

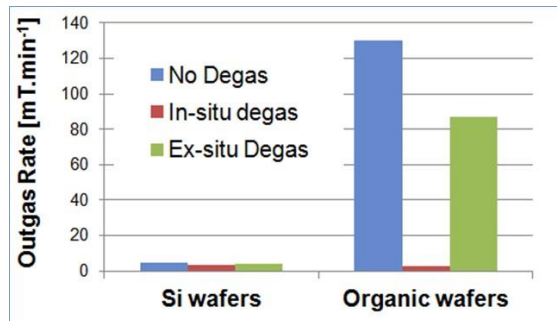


Figure 3. Outgassing Rates of Different Degas Approaches Compared.

It should be noted that a multi-wafer degas approach is only feasible for systems based on pick/place robot wafer handling. Handlers based on a carousel approach where wafers are supported by the handling system throughout the entire process flow are prevented from loading wafers into a multi-slot chamber by the fundamental nature of their design. In those circumstances multi-slot degas hardware can only be located in the atmospheric side of the system, and suffer the problems with re-absorption indicated by the data in Figure 3. Degas at atmosphere may also produce relatively poor temperature uniformity, and risk of particulates if high flow purge gases are used in an attempt to improve temperature uniformity.

In contrast, the SPTS MWD degasses wafers under high vacuum and uses radiative heat transfer producing excellent WIW temperature uniformity ($< 5^{\circ}\text{C}$), low particle levels, and no risk of moisture re-absorption or cross-contamination.

Utmost Integrity

After successful degas, but prior to metal deposition, the FO wafer is pre-cleaned in a plasma etch module. This facilitates the removal of trace oxide layers from the contacts, but due to the composition of the organic dielectric surrounding the contacts, will result in carbon-based material build-up on the chamber walls as organic material is sputtered off the wafer surface alongside the native oxide. This material does not adhere well to ceramic chamber surfaces, and if not carefully managed, can result in early particle failure.

New in-situ paste technologies allow these carbon deposits to better adhere to chamber surfaces during the pre-cleaning process, enabling preventative maintenance intervals that exceed 6,000 wafers. This approach can significantly improve productivity by reducing the frequency of dedicated wafer pastes, which typically require production to be paused every 10 to 25 wafers for chamber pasting when using conventional techniques. The SE-LTX Pre-Clean Etch module from SPTS is designed to increase the pasting interval by $>10\times$ in comparison, resulting in less disruption for production.

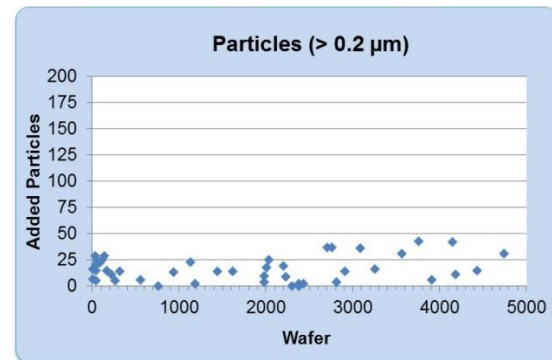


Figure 4. SE-LTX particles maintained at low levels

Exposed Cu Contacts

With the increase in density of interconnects as more complex die such AP and Memory are packaged using FOWLP techniques, there is a need for multiple levels of RDL to overcome interconnect congestion. At the upper most levels $>70\%$ of the wafer surface will be exposed metal contacts, areas much larger than those encountered in mainstream Si UBM/RDL (typically $< 20\%$).

For PVD systems with pre-clean modules based on ICP technology with ceramic chamber walls this increase in exposed metal contact area presents a problem for RF stability. The pre-clean process will involve an 'over-etch' to accommodate any cross-wafer variation in the etch process. During this stage metal contacts on certain parts of the wafer will gradually be revealed as native oxide is removed. Exposed metal will then be sputtered onto the ICP ceramic sidewalls as etching continues to clear all parts of the wafer of native oxide. As more metal coats up the chamber walls the RF coupling through the ceramic to the plasma will deteriorate with the process rapidly falling out of control (Figure 5 shows an accelerated failure case, using wafers with 60% Pure Cu and no native oxide to remove).

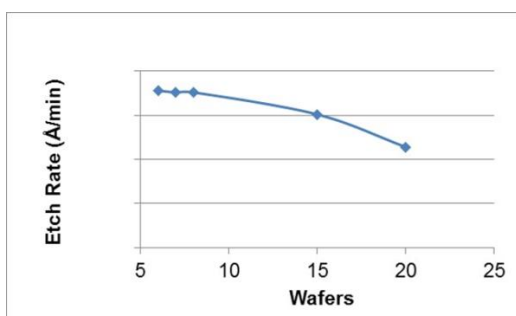


Figure 5. Etch Rate Deterioration Sputtering 60% Pure Cu (native oxide completely removed)

By inserting a discontinuous metal liner in the ICP chamber between the wafer and the ceramic chamber wall it is possible to maintain the ICP effect and sputter metal onto the sidewalls - the design of liner and RF antenna circuit ensures that a continuous band of metal is prevented from forming. With the liner installed Etch Rate stability can be maintained at levels previously experienced with low exposed Cu contact schemes (see Figure 6).

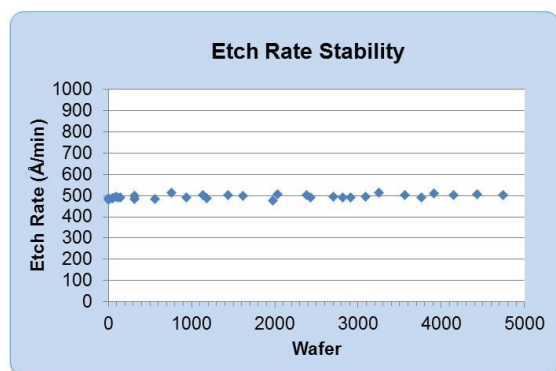


Figure 6. SE-LTX Pre-Clean Etch Stability

Rc Management During Pre-Clean

The use of organic passivation also introduces challenges for Rc that intensify as the CDs of the features reduce. As the surface of the wafer is etched, native oxide is removed from the exposed metal contact, but alongside, the organic passivation is also physically bombarded. This ion bombardment damages the surface of the passivation, releasing C-based volatiles, which in turn, re-contaminates the metal contacts (refer to Figure 7).

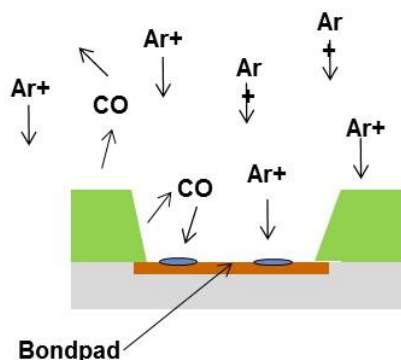


Figure 7. Contamination of Pads During Pre-Clean

Common approaches to tackle this problem include reducing the wafer temperature during process, reducing the volatility of the C by-products. Chilled pedestals and chamber furniture are commonplace as a result. In addition to reducing contaminant volatility vacuum pumping speed is increased to pump away contaminants before they have an opportunity to react with the exposed metal contacts. The use of in-situ pasting during etch also helps with removal of these contaminants, producing stable Rc performance through a batch without drift caused by a back-up of contaminants.

Why 'Soft' ICP Etch is best

Given the challenges of large exposed metal and organic material adhesion with ICP ceramic walls it would be tempting to employ a Charge Coupled Plasma (CCP) or 'Diode Etch' approach to pre-clean, as the single source concept in the form of a RF-driven platen allows the chamber walls to be of metal construction. But the diode etch approach has several disadvantages. First, the diode etch concept, being single source, requires relatively high bias voltages to initiate and sustain a plasma discharge (refer to Figure 8).

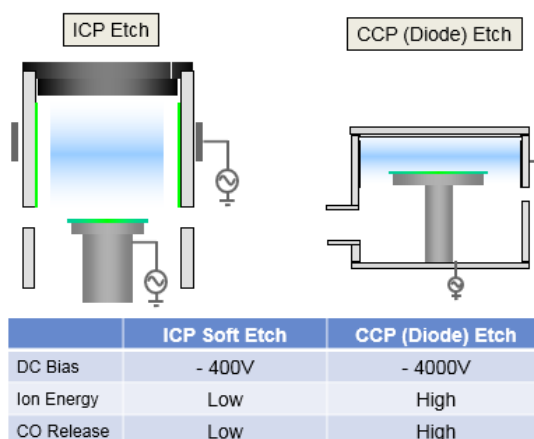


Figure 8. ICP vs Diode Etch Comparison

Where an ICP etch generates wafer biases in the < 500V range, a diode etch module typically works in bias ranges > 3000V. The higher ion bombardment energy involved produces significantly higher levels of contaminant release from the organic passivation on the wafer. For an equivalent maximum wafer temperature during etch, CO partial pressures can be an order of magnitude higher in a diode etch system compared with ICP. This has obvious implications for Rc. The high ion energy will also cause a temperature increase that will need cool steps to curtail given the thermal budget restrictions of mold wafer processing. The cool steps will in turn impact module throughput. With mainstream Silicon UM/RDL these problems can be overcome with the use of ESC clamping combined with backside cooling gas to maintain wafer temperatures at low levels, minimising contaminant volatility. With mold wafer processing however, ESCs are not compatible because of the excessive wafer bows

involved. Another disadvantage of diode etch is that etch rate is typically 50% of an ICP equivalent, limiting the throughput of the system after the degas bottleneck is addressed.

Finally, the diode etch process is less directional than ICP due to its single source and higher discharge pressures utilised. As CDs reduce the aspect ratio of contact topography will inevitably increase. For a diode configuration this will mean longer etches are needed to remove native oxide in the base of the contact. Conversely, the ICP etch, being dual source and low pressure (< 2 mT) produces an anisotropic etch, directional in nature. As aspect ratios increase, the ICP etch is less affected, and throughputs can be maintained.

Wafer Warpage

The benefits described in this paper are not readily attainable, however, unless we can overcome the attendant warping challenges. Epoxy mold wafers can be warped after curing, and the size and shape of the warpage hinge on the different shapes, densities and placement of the die. A FOWLP PVD system must therefore be able to minimize temperature-induced shape shifting, and accommodate wafers with up to 10 mm bow in some cases. The majority threshold for bow is probably lower than 6mm on average however (at least for high density packages), as it's not easy to make uniformly thick, fine resolution conductors on a substrate exhibiting high warpages. This is one of the major challenges facing those seeking to build FOWLP packages on a large area panel format.

Although mold wafers involving carriers typically have bows in the 2-3 mm range, for 'die first' wafers that are unsupported by a rigid carrier it is not uncommon to encounter wafers with 7-8 mm warpage in production (refer to Fig. 9).



Figure 9. Highly Bowed Mold Wafer in EFEM

To successfully handle substrates with warpage of this amount, several aspects of the system design need to be modified. At the atmospheric front end, standard 25-slot FOUPs are replaced with 13-Slot versions. This enables increased clearances for wafer transfer in and out of FOUPs. Clearances at the EFEM Aligner and Transport Module Slot

Valves are also increased. Robot arm velocities are adjusted to prevent wafer movement and processes are tuned to minimise additional warpage generation caused by rapid temperature change. The Sigma[®] fxP system is designed to cope with wafer bows up to 10 mm and its performance has been validated at multiple FOWLP manufacturing sites. One example OSAT has processed > 1 M mold wafers using Sigma and reports wafer breakage rates > 1 in 30k, matching performance of standard silicon processing tools.

High Uptime, Low Cost PVD

Although the bulk of the challenges for volume production are related to wafer handling and pre-treatment there is still room for the metal deposition stage itself to play a part in a high productivity solution. RDL seed depositions are relatively simple and for that reason it's important to use hardware that doesn't exceed requirements. Traditional FEOL PVD chambers designed to offer features such as high deposition rates, directionality, and advanced uniformity control are not so critical for FOWLP. A basic conventional PVD module will meet requirements. For this reason SPTS developed the Inspira PVD module (refer to Figure 10). Inspira PVD has been specifically designed for BEOL processing with all the necessary features a UBM/RDL customer requires.

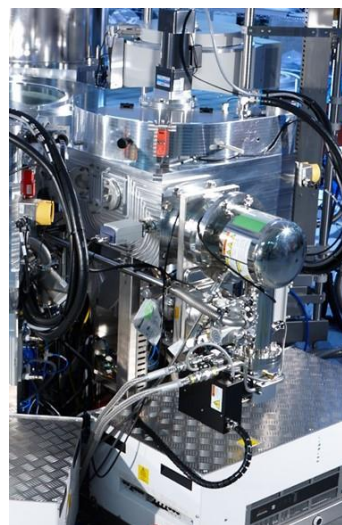


Figure 10. Inspira PVD for Advanced Packaging

In addition to a tailored design, Inspira PVD chambers take advantage of extra thick target compatibility, minimising downtime for PM activity, freeing up engineers to work on other systems. As an example, Inspira Cu modules will run 40% longer than the nearest competitor before requiring maintenance (refer to Fig.11).

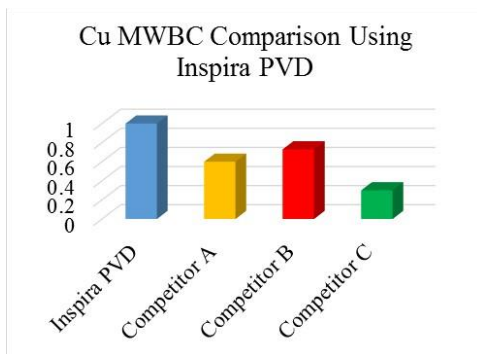


Figure 11. Normalised Cu MWBC comparison.

CONCLUSIONS

The myriad of benefits that FOWLP promises for the production of ultra compact, high I/O electronic devices far outweigh the aforementioned technical barriers to mainstream FOWLP adoption. With the ability to overcome the contamination, particle and wafer warpage challenges that can impede FOWLP implementations, electronics manufacturers can unlock the full potential of FOWLP while eliminating frictions affecting production speeds and yields.

REFERENCES

- [1] Azémar J, Buisson T “Fan-Out WLP Technology: Status, Market & Trends” Yole Developpment, SPTS Packaging Symposium, March 16, 2016