

Lithography Technique to Reduce the Alignment Errors from Die Placement in Fan-out Wafer Level Packaging Applications

Warren Flack, Robert Hsieh, Gareth Kenyon, Khiem Nguyen, Manish Ranjan
Ultratech Inc, 3050 Zanker Road, San Jose. CA. 95124
wflack@ultratech.com 1-408-321-8835

Nuno Silva, Paulo Cardoso, Eoin O Toole
Nanium S.A., Avenida 1 de Maio 801, 4485-629 Vila do Conde, Portugal
Eoin.OTOole@nanium.com 351-252-246-000

Rainer Leuschner, Werner Robl, Thorsten Meyer
Infineon Technology AG, Wernerwerkstr. 2, 93049 Regensburg, Germany.
Rainer.leuschner@infineon.com 49-941-202-3802

Abstract

The rapid growth of wireless consumer electronics products is driving demand for cost effective and small form factor packaging solutions. While front end silicon technologies have followed Moore's law by device scaling, the back end infrastructure has lagged in similar advancements. This has created an interconnect gap whereby the signal speed achieved on silicon is significantly higher than the speed achieved on the printed circuit boards. Innovative advancements such as Fan-out wafer level packaging technology were introduced to address the pad limitation consideration with traditional wafer level packaging while delivering miniaturization and potential low cost packaging advantages. It does this by extending the package interconnect area beyond the front end chip size to allow increased number of I/O required for large die sizes. This technology allows tested-good dice to be reconstituted into wafer form, and interconnections are formed using wafer level processing technology.

Die positioning control within the reconstituted wafer significantly affects downstream process requirements. The use of high productivity pick and place equipment with multiple gantries create challenges for the lithographic tool alignment when die placements from each gantry are not identical. This will be especially true in the future as the placement tolerances are reduced for advanced products containing multiple die types.

This paper describes the inaccuracy in pick and place from single and dual gantry operation, and investigates lithographic alignment methods specifically developed to minimize pick and place errors from multiple gantry operation. The current single zone alignment algorithm was extended to create multiple selection zones to match the multiple gantries of the die pick and place equipment. The enhanced capability allows the flexibility to conduct a separate alignment mapping for different zones of the reconstituted Fan-out wafers. The dual zone mapping gave more effective compensation for a gantry matching error, resulting in better than 50% improvement in registration error compared with a single zone mapping. This provides significantly superior alignment control for next generation devices fabricated with fan out wafer level packaging process.

Introduction

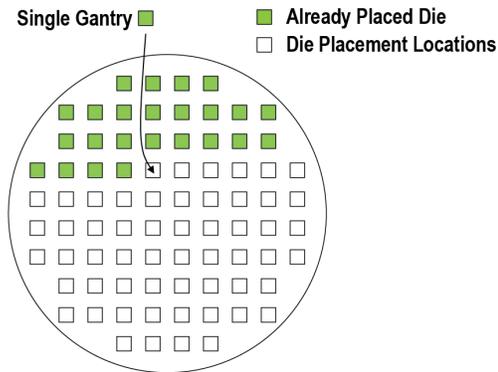
Leading edge packaging technologies are expected to play a significant role in managing the transition between rapidly shrinking silicon features and slowly shrinking second level interconnect dimensions. Fan-out wafer level packaging technology is gaining significant interest from major cell phone manufacturers as a potential cost effective packaging solution. [1]

During the Fan-out wafer level packaging process, the silicon chips are probed, thinned and singulated. These chips are then placed on an adhesive tape carrier and subsequently placed on a sacrificial metal carrier using standard pick and place equipment. The next process sequence involves compression molding for embedding the silicon chips. After post mold curing the adhesive tape and reconfigured wafer are released from the metal carrier. Once the panel fabrication is completed, thin-film technologies such as sputtering, photolithography and electroplating are performed. The final interconnect process step is the ball drop process followed by singulation and testing of packages. [2]

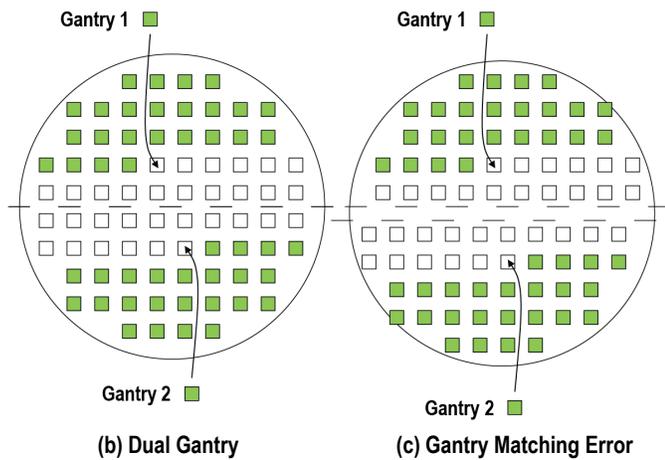
As Fan-out technology transitions into high volume manufacturing an optimum balance is sought between performance and cost. Unlike manufacturing for the front end, which is largely performance driven, the capabilities of existing tools provide options regarding error budgets and levels of sophistication required at each step in the manufacturing process. For example, to allow for typical variation in the Fan-out panel creation process, wafer bow allowance is greater than typically seen in front end semiconductor manufacturing. This tradeoff places additional requirements on downstream processing equipment.

One important performance issue for Fan-out is the placement of die on the substrate. The two operations that most affect die location error are the die placement operation and die migration during the compression molding process. Die pick and place on a wafer is shown in Figure 1 by filled rectangles for die already placed, and outlines of rectangles for subsequent die placements. The task of die placement is fairly straightforward using a single gantry mode of operation, as shown in Figure 1(a). However for increased productivity a dual gantry configuration is needed, as shown in figure 1(b).

For single gantry mode, a positional offset of the gantry affects all die placements equally, and does not affect the placement precision within the set of die. However for a dual gantry configuration, positional offset of any gantry can produce a matching error between gantries. Figure 1(c) illustrates a translational matching error between two gantries.



(a) Single Gantry



(b) Dual Gantry

(c) Gantry Matching Error

Figure 1: Single gantry (a) and dual gantry (b) methods can be used for die placement. The dual gantry method of die placement populates the wafer faster than the single gantry method, however the dual gantry configuration also introduces possible gantry matching error as shown in (c).

This creates significant challenges in the photolithography process steps for aligning subsequent vias and metal layers to the device pattern. The die migration during compression molding is a second major source of placement error. However this effect is more random in character and thus more difficult to address using real time metrology. [3] This paper investigates registration issues from the perspective of stepper lithography, and demonstrates an effective method for handling general dual gantry matching errors automatically in the run time stepper operation.

During a conventional stepper alignment routine, all linear grid terms can be estimated by calculating a least squares fit of the measured error as a function of the field location. Using Enhanced Global Alignment (EGA) additional sites can be added to the alignment sampling to minimize effects of individual measurement error. [4] However this alignment approach was designed for conventional front end silicon

processes that can be effectively optimized by considering a single full wafer zone. This approach is not optimal if the wafer has multiple zones, each having different systematic die placement errors.

Since estimation of higher order registration error components from sparse measurement data is subject to error, it is desirable to retain a linear model when possible and minimize contributions from higher order error sources. [5,6] The error arising from dual gantry matching can be simplified by splitting the problem into two linear optimization zones, one for each gantry. Figure 2(a) illustrates this using an example layout where the gantry 1 zone (upper half) is shifted in X and Y with respect to the gantry 2 zone (lower half). Figure 2(b) shows that a linear optimization with a single zone (red) linear fit to the entire data set provides a partial optimization, and Figure 2(c) shows that a better fit can be achieved by splitting the data into two zones that match the upper gantry (red) and the lower gantry (blue) and performing separate linear optimizations.

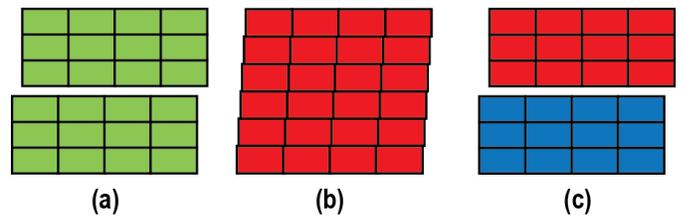


Figure 2: Dual gantry zone error (a) represented as (b) single zone linear fit and (c) a dual zone linear fit. The dual zone model can accurately represent an abrupt shift between gantry zones whereas the single zone model linearly distributes the error across the field.

Experimental Methods

A dual-gantry system, 8800 CHAMEO from Datacon, was used to construct a lot of twelve 300mm Fan-out product wafers. The 8800 system was selected for its high-speed and high-accuracy performance. As described in the previous section, each gantry of the dual-gantry system would define a unique zone with a fixed array offset.

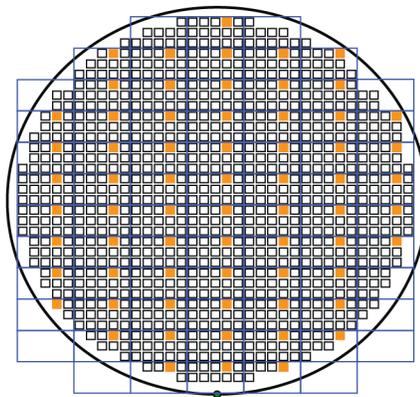
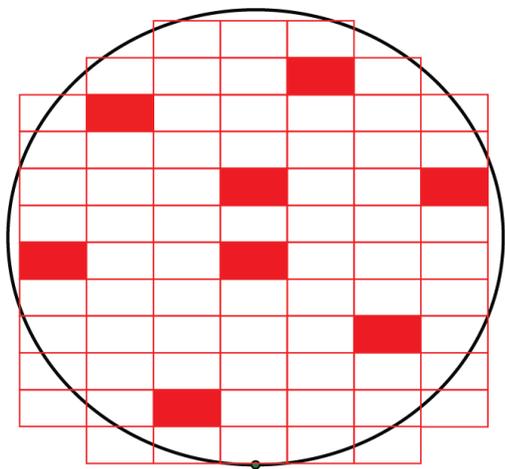


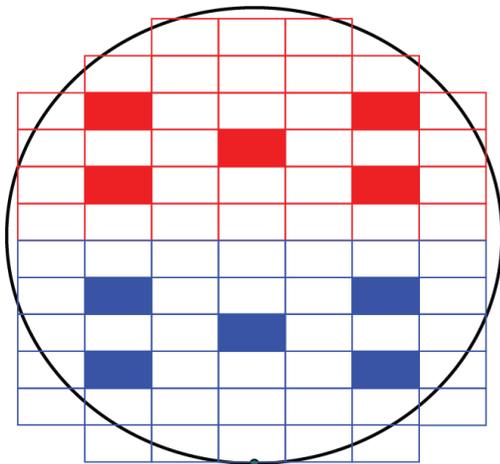
Figure 3: Wafer map showing lithography stepper fields, each consisting of a 5x3 die array. Within each stepper field (outlined in blue) is a metrology die (orange). All vector maps reflect the use of this metrology die position within the field.

In this investigation, the top-half zone of the wafer corresponding to gantry 1 was programmed with a $30\mu\text{m}$ X and Y array shift relative to the bottom-half zone of the wafer corresponding to gantry 2. The 300mm wafer layout and die array are illustrated in Figure 3. Each lithography stepper field covers a 5 column by 3 row die array with the metrology die shown in orange.

Lithography for this investigation was performed using an Ultratech Unity AP300 Wafer Stepper. This system has a 0.16 numerical aperture (NA) and employs broadband Mercury ghi-line illumination from 350 to 450 nm wavelength. The low NA and broadband illumination spectrum of the stepper provides a large depth of focus, which is a requirement for Fan-out applications where the wafer topography and variation of the die height are much greater than the standard process on silicon wafers. [7] The stepper was equipped with an Advanced Warped Wafer Handling option (AWWH) to handle the highly warped substrates.



(a) Single Gantry Zone Mapping



(b) Dual Gantry Zone Mapping

Figure 4: Wafers drawings showing (a) single gantry zone mapping sites (red) and (b) dual gantry zone mapping sites (gantry 1 red and gantry 2 blue).

Six 300mm Fan-out wafers were aligned with single zone mapping and exposed. The single zone mapping layout has

seven mapping sites on a wafer as a whole. The other six 300mm Fan-out wafers were aligned with dual-zone mapping and exposed. The dual-zone mapping layout divides the wafer layout into two independent zones with five mapping sites per zone. Figure 4 compares the wafer layout of the single-zone mapping versus the dual-zone mapping. Note that this is representative of typical sampling plans and the number and location of mapping sites can be modified in the stepper process job as desired. [7]

Although this investigation specifically studies dual zone mapping with a simple translational error between zones, the dual zone mapping can correct for other linear grid errors within each zone such as rotation and scaling. The method can also extend to more than two zones if needed.

Alignment on the stepper is performed using a Machine Vision System (MVS) alignment system. [8] With MVS the stepper uses pattern recognition techniques to align to any unique feature in a die in the stepper field. In this investigation, the large cross at the die corner was trained as a target for alignment as shown in Figure 5. The smaller cross directly below the large cross is the reticle reference mark used for through the lens alignment.

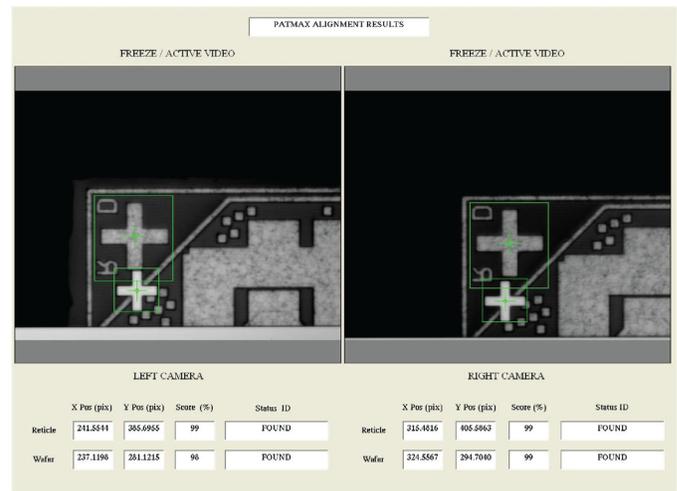


Figure 5: Machine Vision System (MVS) view of wafer targets (large cross) and reticle reference mark (lower small cross) for left and right side of the stepper field of view.

The photoresist material used in this investigation was a standard commercially available polyimide spun on at a thickness of $7\mu\text{m}$. The developer is a standard commercially available metal ion free chemistry. Pre- and post- baking were performed at 110°C . A 2 hour wait time was used between the bake and exposure steps.

The registration metrology was done on a Nikon NEXIV VMR-H3030 automated video measuring system. The tool features a $300 \times 300 \times 150\text{mm}$ stage, TTL laser height/profile scanning and intelligent pattern recognition. It is capable of accuracy and precision down to $0.01\mu\text{m}$. The registration error was measured as shown in Figure 6. The error is the offset between the center of the metal pad and the patterned polyimide dielectric opening.

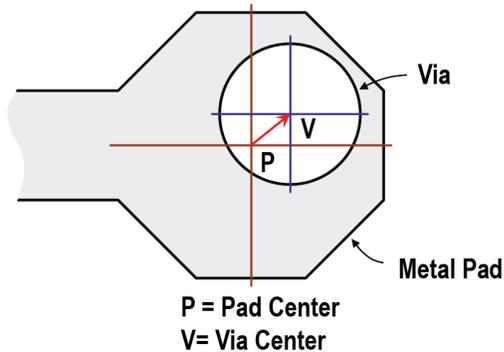


Figure 6: Registration is measured as offset (red arrow) between the center of dielectric opening (via) and center of the metal pad.

Results and Discussion

Figure 7 illustrates the registration error from using a single zone optimization for a dual zone error. This example shows a translation shift, G (purple), in X and Y between the gantry 1 zone (top) and the gantry 2 zone (bottom). This is an example of a gantry matching error.

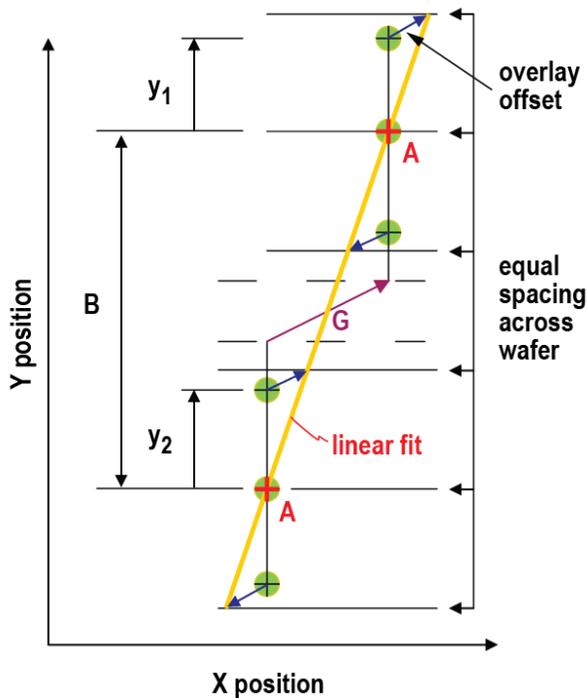


Figure 7: Schematic diagram showing effect of translational matching error, G, between gantries. Modeled linear solution (yellow) provides best match at the alignment marks (A), however residual error is proportional to Y distance from the effective alignment site within each respective zone. The registration error (blue) is proportional to the Y distance from the effective alignment site in each zone.

Assume that each zone is perfectly linear. The linear fit model (yellow) from the stepper cannot match the non-linear pattern at all locations, however the best fit is achieved at the effective alignment locations, A (red), separated by a

distance, B. From this graphical construction one can see that the registration error (blue) is proportional to the Y distance from the effective alignment site in each zone. For example in the gantry 1 zone the error is proportional to y_1 , and in the gantry 2 zone the error is proportional to y_2 . Therefore the registration error (X_{offset} , Y_{offset}) is a function of the Y distance from the alignment mark locations:

For gantry zone 1:	For gantry zone 2:
$X_{offset_1} = y_1(G_x/B)$	$X_{offset_2} = y_2(G_x/B)$
$Y_{offset_1} = y_1(G_y/B)$	$Y_{offset_2} = y_2(G_y/B)$

where G_x and G_y are the X and Y components of the translational shift G (purple).

For the same dual gantry wafer, Figure 8 illustrates that the registration error is more effectively reduced by using a dual zone mapping. Since each zone is independently modeled, the shift, G (purple), between zones can be directly accounted for. The remaining registration errors (not shown in figure 8) are from non-linear and random errors in each zone.

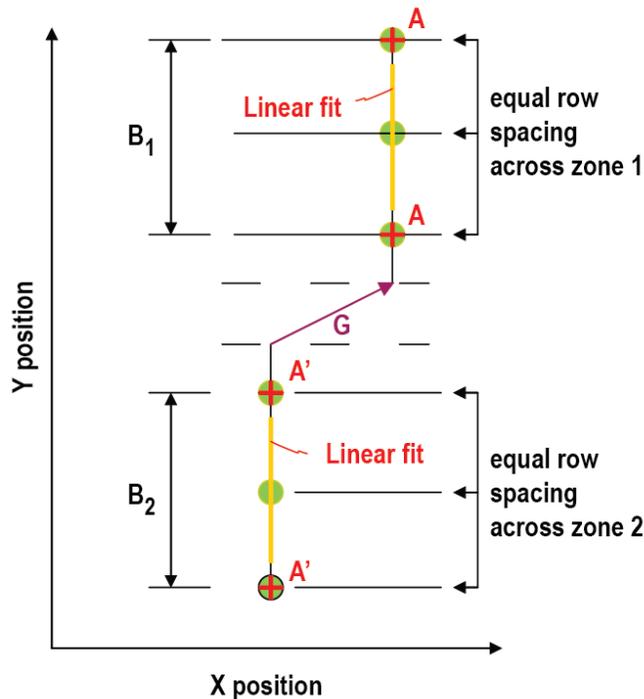
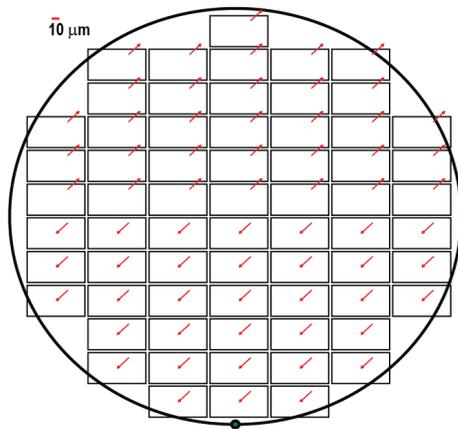
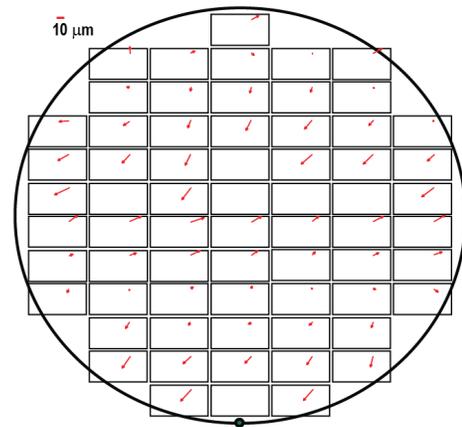


Figure 8: Schematic diagram showing how dual zone mapping can effectively account for the dual gantry matching error, G. EGA linear fit within each zone can then be more effective in reducing random errors within each zone.

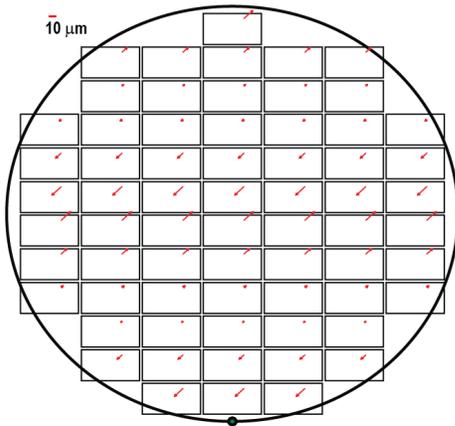
An intentional offset of 30 μm in X and Y between top and bottom gantry zones, as shown in Figure 9(a), was used to evaluate the effectiveness of stepper software for single and dual mapping for controlling dual gantry registration error. A modeled registration error map for single zone mapping of a dual gantry wafer with intentional error of 30 μm in X and Y is shown in figure 9(b). The modeled error from dual zone mapping is not shown in this comparison since all registration vectors would be zero in the absence of non-linear and random errors.



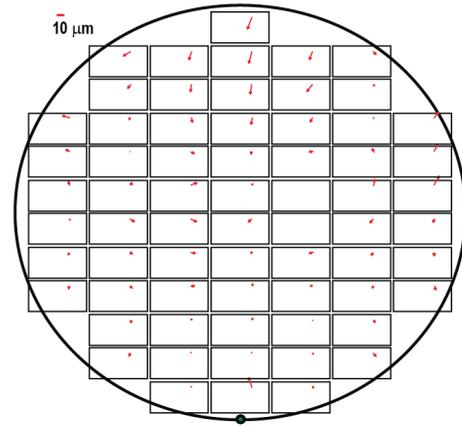
(a) Programmed Error



(a) Single Zone Mapping



(b) Model Single Zone Mapping



(b) Dual Zone Mapping

Figure 9: Modeled registration error for (a) Intentional gantry error with translation of $30\ \mu\text{m}$ in X and Y between top and bottom zones while (b) exhibits the modeled registration result for single zone mapping. The vector locations match the metrology die position within the stepper field.

Figure 10: Experimental registration data for (a) dual gantry, single zone mapping and (b) dual gantry, dual zone mapping. Single vectors at each measurement site represent the average of six wafers in each set, and vector locations match the metrology die position within the stepper field.

The large $30\ \mu\text{m}$ in X and Y offset between top and bottom gantry zones created data collection problems for metrology. At large offsets, the metrology structure became difficult to read, and several uncharacteristically large fliers removed from the data set were attributed to this problem. For each mapping mode, six wafers were measured and averaged to produce a single error vector at each measurement site.

Figure 10(a) shows experimental registration data using single zone mapping on a wafer set with an intentional gantry translational error (G) of $30\ \mu\text{m}$ in X and $30\ \mu\text{m}$ in Y. Figure 10(b) shows experimental raw data using dual zone mapping software on a wafer set with the same intentional gantry error. The dual zone mapping result shows better registration performance than the single zone mapping result. Note that the experimental single zone error signature in Figure 10(a) resembles the modeled error in Figure 9(b).

Figure 11 shows the same data as in Figure 10 plotted in histogram form for X and Y. Superimposed on the histograms are calculated normal distribution curves based on the mean and standard deviation of the registration measurements. The

Dual gantry zone mapping produces a tighter, more Gaussian distribution than the single gantry zone mapping. The long distribution tails for the dual zone Y data may indicate an additional error source. A statistical summary of the experimental registration results are shown in Table 1. The dual gantry zone mapping shows X and Y three sigma values are 51% and 73% of the single gantry zone mapping. The dual zone mapping provides more effective compensation for a gantry matching error, resulting in a much lower registration error than is possible with a single zone mapping. The experimental results show a good match to the modeled data, providing a foundation for further error investigations.

3 sigma (μm)	X	Y
Single Gantry Zone Mapping	26.3	23.8
Dual Gantry Zone Mapping	13.6	17.7

Table 1: Dual gantry registration errors using dual and single zone mapping. Dual zone mapping significantly reduces the registration error from gantry matching error.

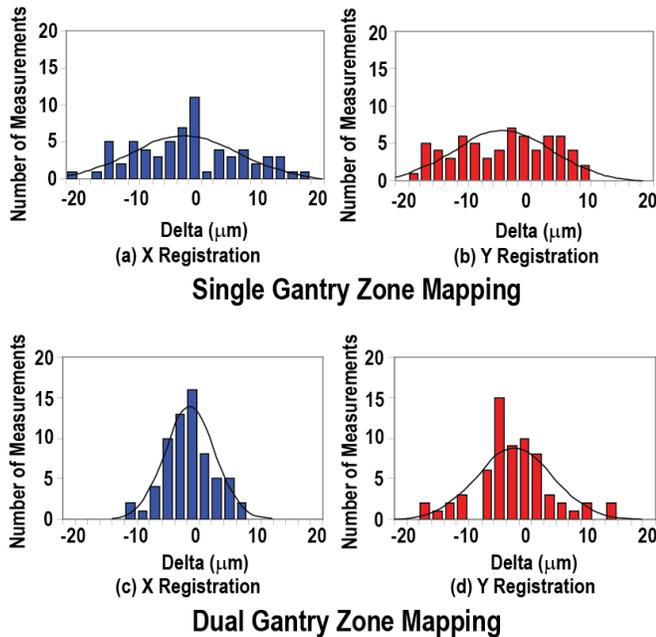


Figure 11: Registration histograms for X and Y for single gantry zone mapping (a) and (b); and dual gantry zone mapping (c) and (d). Superimposed on each histogram is a calculated normal distribution curve matching the mean and standard deviation of the particular measurement set.

The effect of dual zone mapping on lithographic productivity is minimal. In automatic operation, the time required for single zone and dual zone mapping operations is similar and primarily depends on the number of alignment sites required for each method. For wafers that are highly linear, the single zone mapping would be slightly faster, however this advantage decreases if single zone mapping sites are added to maintain registration requirements in the presence of a gantry matching error. For this experiment, measured lot throughput was comparable, within 1%.

Conclusions

The current single zone alignment algorithm was extended to create multiple selection zones to match the die pick and place gantry zones within the wafer. The enhanced capability allows the flexibility to conduct a separate alignment mapping for different zones of the reconstituted wafers. The dual zone mapping gave more effective compensation for a gantry matching error, resulting in better than 50% improvement in registration error compared with a single zone mapping. This provides significantly superior alignment control for next generation devices fabricated with fan out wafer level packaging process. The experimental data verified that dual mapping addresses the additional error source in going from single to dual gantry configuration. The effect of dual zone mapping on lithographic productivity was measured and found to be minimal. In addition to handling a translational shift between gantries, the method also compensates for linear grid errors within each zone, and can be extended to more than two zones.

The large intentional gantry zone offset for this investigation created data collection problems for the metrology equipment. Therefore it is likely that metrology

error comprises a significant portion of the residual error. For future work the intentional metrology error should be reduced to minimize metrology problems. Further investigations are also needed to understand and minimize the remaining error components such as die migration in the compression molding operation.

Acknowledgments

We would like to thank Alberto Martins of Nanium for preparation of the dual gantry die placement job with an intentional array offset. We would also like to thank Nelson Pinho of Nanium for his assistance with registration metrology.

References

1. Johnson, S., "STATS ChipPac to Manufacture Infineon's First-Gen eWLB Products", *Semiconductor International*, August 2008.
2. Brunnbauer, M. *et al*, "Embedded Wafer Level Ball Grid Array (eWLB)", *Electronics Packaging Technology Conference 8th Proceeding*, December 2006.
3. Chee Houe Khong, *et al*, "A Novel Method to Predict Die Shift During Compression Molding in Embedded Wafer Level Package", *Electronic Components and Technology Conference Proceedings*, May 2009.
4. Hsieh, R. *et al*, "Lithography Challenges and Considerations For Emerging Fan-Out Wafer Level Packaging Applications", *International Wafer-Level Packaging Conference Proceedings*, October 2009.
5. Perkins, M. *et al*, "Intermix technology: The Key to Optimal Stepper Productivity and Cost Efficiency", *Optical/Laser Microlithography V Proceedings*, SPIE 1674 (1992).
6. Van den Brink, M. *et al*, "Performance of a Wafer Stepper with Automatic Intra-die Registration Correction", *Optical Microlithography VI Proceedings*, SPIE 772 (1987).
7. Todd, B. *et al*, "Thick Photoresist Imaging Using a Three Wavelength Exposure Stepper", *Micromachining and Microfabrication Process Proceedings*, SPIE 3874 (1999).
8. Flack, W. *et al*, "Application of Pattern Recognition in Mix and Match Lithography", *Optical/Laser Microlithography VIII Proceedings*, SPIE 2440 (1995).