Extended Abstracts of the 2020 International Conference on Solid State Devices and Materials, VIRTUAL conference, 2020, pp147-148

Die-Level Cu-CMP Technology in Via-Last TSV Process for Multichip-to-Wafer 3D integration

Shuai Liu[1], Kousei Kumahara[2], Yuki Miwa[2], Hisashi Kino[3], Takafumi Fukushima[1,2], and Tetsu Tanaka[1,2]

[1] Dept. of Mechanical Systems Engineering, Graduate School of Engineering, Tohoku University,
6-6-12 Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan
Phone: +81-22-795-6978, E-mail: link@lbc.mech.tohoku.ac.jp
[2] Dept. of Biomedical Engineering, Graduate School of Biomedical Engineering, Tohoku University,
[3] Frontier Research Institute for Interdisciplinary Sciences, Tohoku University

Abstract

Cu-CMP is an essential process to polish off thick Cu overburden for fine-pitch redistribution layer (RDL) formation on through silicon via (TSV) in 3D integrated circuits (3D-IC). In this paper, die-level 3D integration with via-last TSV formation is studied to three-dimensionally stack incoming known good dies (KGDs) without TSV. Wafer-level Cu-CMP is industrially used, but die-level Cu-CMP is very challenging. We investigate the several parameters of die-level Cu-CMP technology by conditioning pad rigidity, die arrangement, and polishing rate with enhanced chemical effects. Chipping defects are perfectly eliminated by optimizing the Cu-CMP conditions, and the in-plain uniformity of the Cu-CMP process is well controlled.

1. Introduction

TSV formation is a core technology of 3D-IC and mainly divided into two categories, via-middle and via-last TSVs. The former TSVs are fabricated only by wafer-level 3D integration. And the latter TSVs can be applied to wafer- and die-level 3D integration because TSV formation is required for completed 2D chips and wafers fabricated in shuttle foundry service without TSVs. The existence of huge numbers of TSVs allows effective vertical interconnection between chips. Thus, inter-chip communication can be much improved by the massively parallel interconnection to significantly reduce the latency and drastically increase the bandwidth [1]. Electroplated Cu is often used as a conductive material for TSVs because of its excellent electrical properties, cost-effectiveness, and low-temperature methodology. Therefore, the Cu-CMP technology is employed in the TSV formation process to polish off the overburden caused by Cu filling. The die-level Cu-CMP optimization discussed in this paper is focusing on chip-to-wafer and multichip-to-wafer 3D integration we have developed so far [2].

2. Die-Level 3D Integration Based on Via-Last TSV

A process flow of die-level 3D integration with via-last TSV formation is described in Fig. 1. Incoming KGDs fabricated in the shuttle service of chip foundries are temporarily bonded to a carrier wafer in a face-down fashion. Then, the dies are thinned down to 50 μm or less by grinding, followed by Si-CMP. After that, deep Si via holes are formed on the backside of the KGDs by Bosch etch. After a dielectric SiO_2 liner is deposited on the holes by plasma CVD, the bottom dielectric liner in the deep via holes is selectively etched by RIE. The subsequent process is barrier and seed layer deposition by sputtering, followed by electroplating to fulfill the deep Si via holes with Cu. There are two options to remove the Cu overburden on the resulting Cu-TSVs: one is wet etching [3], and the other is Cu-CMP. The Cu wet etching is technologically possible, but excess side etching is a big concern for high-density RDL formation. To address the interconnection issue, we employ Cu-CMP to remove the overburden. After Cu-CMP, RDLs are formed on the Cu-TSV, and solder microbumps are formed by semi-additive electroplating. The thinned KGDs with the Cu-TSV are three-dimensionally stacked, and then debonded from the carrier wafer, and finally transferred to the corresponding KGD or Si interposer to give 3D-IC. By repeating these processes, additional KGDs are sequentially stacked for fabricating heterogeneous 3D-IC in layers. We have previously fabricated a 3D image sensor [3] and 3D microprocessor chip etc. at the die level.

3. Experimental Methods

50-μm-thick Si dies with a Cu/Ti/SiO_2 (7 μm/100 nm/4 μm) layer in order from the top were fabricated by temporary bonding with a glass wafer, thinning, SiO_2 CVD, Ti/Cu (100 nm/200 nm) sputtering, and Cu electroplating. The 7-μm-thick Cu overburden was removed by Cu-CMP. We used two different polishing pads, a hard polyurethane pad IC1000™ and soft one composed of a double layer of polyurethane and textile, as shown in Fig. 2. We also employed two different slurries, standard and high-speed types. The following Cu-CMP conditions were kept at constant: the rotation speed of the base plate: 15 rpm and CMP pressure: 7.5 kPa.

4. Results and Discussion

We first evaluate the impact of the two types of polishing pads on Cu-CMP. As seen from the top picture of Fig. 3 (a), the edges of dies temporarily bonded to the center of the carrier wafer are excessively polished. The Si layer is exposed using the hard pad IC1000^{TM} through the excess polishing beyond SiO_2 underneath the Ti layer. That causes serious Cu contamination, and electrical short will occur between the neighboring TSVs designed at the peripheral area of the die. In addition, the Cu overburden at the center of this die remains, as shown in the bottom picture of Fig. 3(a). On the other hand, almost all the SiO_2 layer remains using the soft pad after Cu-CMP, as shown in Fig. 3(b). This is due to the die-edge polishing mechanism, as shown in Fig. 2. The soft pad can follow the die shape and give no excess load to the edge of the dies. Furthermore, the inter-die variation with the

soft pad is much lower than that with the hard pad. The following experiments listed in Table I are done with the soft pad. The standard and high-speed slurries are compared with the samples No. 1 and No. 2 in Table I. As a result, it is found that the Cu-CMP rate with the high-speed slurry is 16 times higher than the standard one although the slurry/H_2O_2 ratio is different. In addition, it is worth noting that the number of chipping parts is remarkably reduced by using high-speed slurry. Compared with the standard slurry, the high-speed one has a much higher ratio of organic acid and inorganic salt, but a smaller proportion of abrasive particles. This means that the standard slurry is biased toward mechanical mechanisms, whereas the high-speed slurry is biased toward chemical mechanisms. The effect of die arrangement, types A and B, on Cu-CMP is evaluated as shown in Fig. 4. From the results comparing the sample No. 2 with sample No. 3, the type B arrangement shows high-yield polishing with less chipping (see Table I). However, the CMP rate is lower than the type A arrangement. This is probably due to the amount of slurry supplied to the center die through the path in-between dies. The chipping failures are solved with the composition of a high-speed slurry/H_2O_2 ratio of 90/10 by volume (sample No. 4 in Table I). This is very mild conditions having the Cu-CMP rate of 500 nm/min.

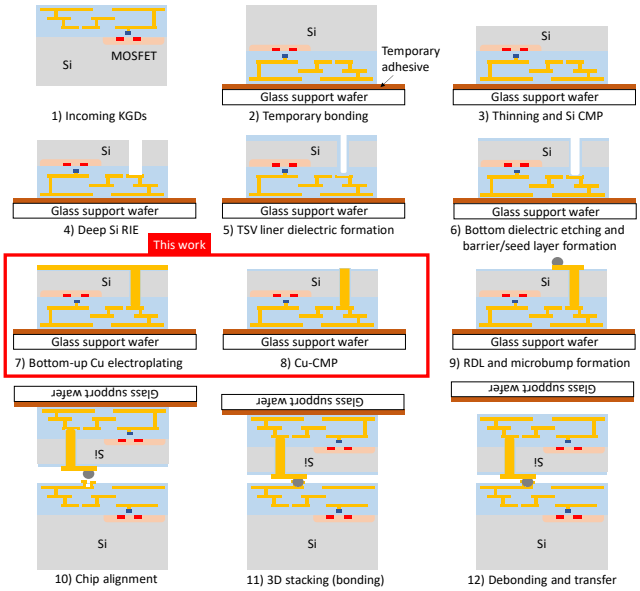

Fig. 1 Die-level 3D integration flow

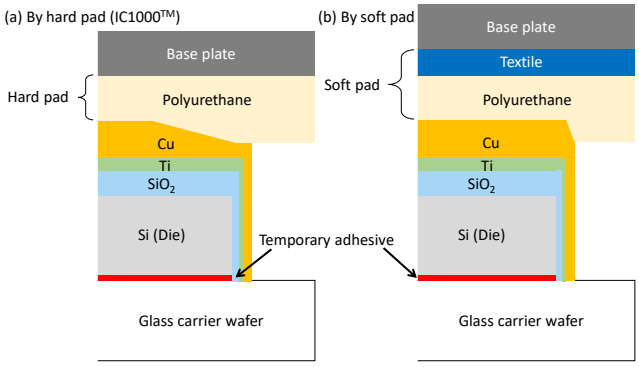

Fig. 2 Polishing mechanism by die-level Cu-CMP: cross-section of die edges polished by hard (a) and soft (b) pads.

5. Conclusions

In this via-last 3D integration study, the die-level Cu-CMP process is evaluated with two types of polishing pads, slurries, and die arrangements. Compared with a hard pad (IC1000™), a soft pad has higher global polishing uniformity and lower chipping probability. The soft pad can also be used for polishing dies with warpage. When the H_2O_2 content in the high-speed slurry is decreased to 10vol%, the die-level Cu-CMP works very well. The high-speed slurry can exhibit a moderate polishing rate and keep the die intact. In addition, the high-density die arrangement B would further increase Cu-CMP yield, compared to another low-density arrangement A, leading to great prospect in future multichip-to-wafer 3D integration with via-last TSV.

Acknowledgment

JSPS KAKENHI Grant-in-Aid for Scientific Research (A), Grant Number:18H04159, supports this work. This work was also supported through the activities of VDEC, The University of Tokyo, in collaboration with Cadence Design Systems.

Reference

[1] M. Koyanagi *et al. Proc. IEEE*, 97 (2009), 49-59.
[2] T. Fukushima *et al. IEDM Tech. Dig.* (2007), 985-988.
[3] K. Lee *et al.*, *IEEE Trans. ED,* 60 (2013), 3842-3848.

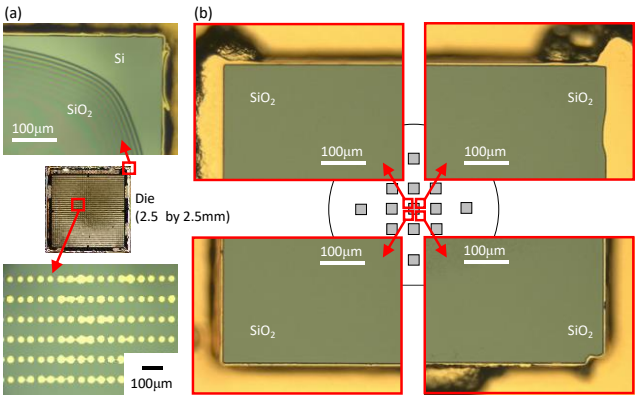

Fig. 3 Cu/SiO_2 selectivity differences between hard (a) and soft (b) pads by die-level Cu-CMP.

Table I Die-level Cu-CMP conditions and results with a soft pad.

No.	Slurry	Die arrange-ment	Time (min)	Polishing rate (nm/min)	Number of chipping
1	Standard: 31% H_2O_2 = 95 : 5	A	50	140	36
2	High-speed: 31% H_2O_2 = 50 : 50	A	3	2300	6
3	High-speed: 31% H_2O_2 = 50 : 50	B	8	880	2
4	High-speed: 31% H_2O_2 = 90 : 10	A	14	500	0

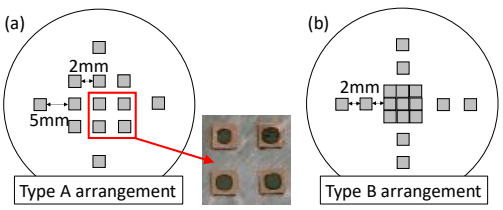

Fig. 4 Die arrangement types A (left) and B (right).