Extended Abstracts of the 2019 International Conference on Solid State Devices and Materials, Nagoya, 2019, pp417-418

Size-Reduction of HBW System using WOW Bumpless TSV Interconnects

Takayuki Ohba[1]

[1] Heterogeneous and Functional Integration Unit, Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama 226-8503, Japan
Phone: +81-45-924-5866 E-mail: ohba.t.ac@m.titech.ac.jp

Abstract

The benefits of bumpless interconnects for reducing the size of 3D systems are described. A bumpless through-Si-via (TSV) process followed by a wafer-on-wafer (WOW) process gives no gap between dies and a higher density than when using bump electrodes. In addition, the physical lengths of the interconnects and the die pitch can be reduced to below 10 μm by ultra-thinning 300 mm wafers.

By reducing TSV pitch, TSVs with a density higher than 10^6/cm^2 can be allocated on a chip surface. Short interconnects formed at high density have much lower impedance (< 1/10) and thermal resistance. In the case of a DRAM multi-level stack, the occupation ratio of I/Os at the device surface becomes small, and heat dissipation occurs through the I/Os.

Therefore, by using a bumpless interconnect process, large memory capacity and bandwidths as high as 8 TB/s or greater can be realized without sacrificing power efficiency and thermal resistance.

1. Introduction

Semiconductor devices and computer systems have evolved as feature sizes have been continuously reduced. According to this trend, system volumes will be 50 mm^{33} and the power consumption will be 0.5 mW [1][2]. Even in such small computers, high performance and large memory capacity are desired without sacrificing power efficiency. Conventional two-dimensional (2D) scaling, however, will be forced to face an economic crisis due to the expensive lithography processes and facilities required [3][4].

A promising approach to overcome these problems is to combine 3D stacking with 2D scaling, i.e. *co-integration* extended into the third dimension (z-direction). In detail, the z-height of a multi-wafer stack must be small, meaning that there should be no bumps between dies, and the dies should be thin. This paper describes size reduction achieved by using bumpless TSV interconnects and ultra-thinning of 300 mm wafers. The technology described here can be used for next-generation high-bandwidth 2.5D and 3D systems.

2. Bumpless Interconnects

Figure 1 shows a comparison of bump and bumpless interconnects using TSVs, assuming eight dies for a memory core and one logic controller. Since a chip-level stack needs a pick-and-place process and bump connections, the die thickness is limited due to the stiffness requirements and warpage, resulting in a chip pitch of around 80–100 μm. Since the WOW process consists of *Bonding First* using a thinned wafer and then *TSV Interconnects*, the wafer thickness is determined by whether thinning degrades the device

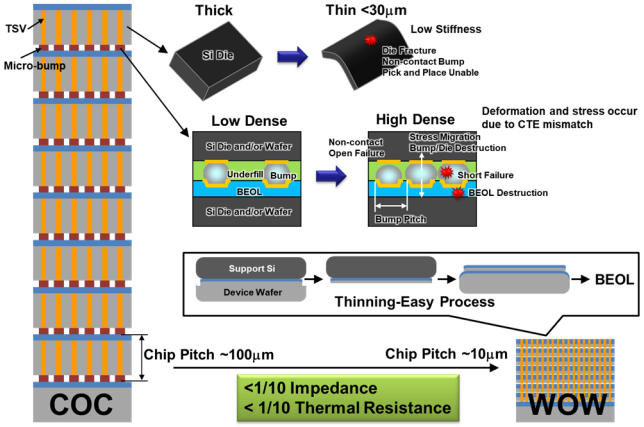

Fig. 1. A comparison of 3D stack structure with and without bump electrodes, where COC and WOW are Chip-on-Chip and Wafer-on-Wafer, respectively.

characteristics. There was no damage when a Si wafer for DRAM was thinned to 4 μm [5]. The wafer (chip) stack pitch becomes around 10 μm, which is 1/10 thinner than that of COC.

3. Short, High-Density TSV Interconnects

The WOW process enables wafer thinning from 775 μm to 1 μm, as shown in Figure 2. Since the physical length of TSV interconnects becomes the wafer thickness, including the device layer and adhesive, the total length in the case of an eight-wafer stack will be <80 μm. Trends of TSV interconnects versus the number of stacked chips and/or wafers are estimated as shown in Figure 3. The total height, based on the die-to-die pitch, is less than 0.5 mm even for a stack of 60

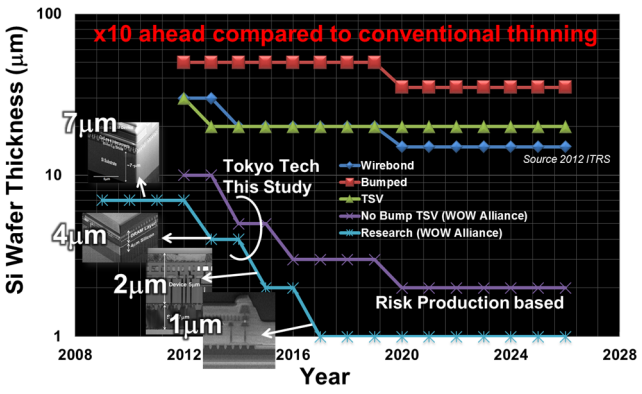

Fig. 2. Trend of 300 mm silicon wafer thickness. Ultra-thinning down to 1 μm thickness has been achieved [6].

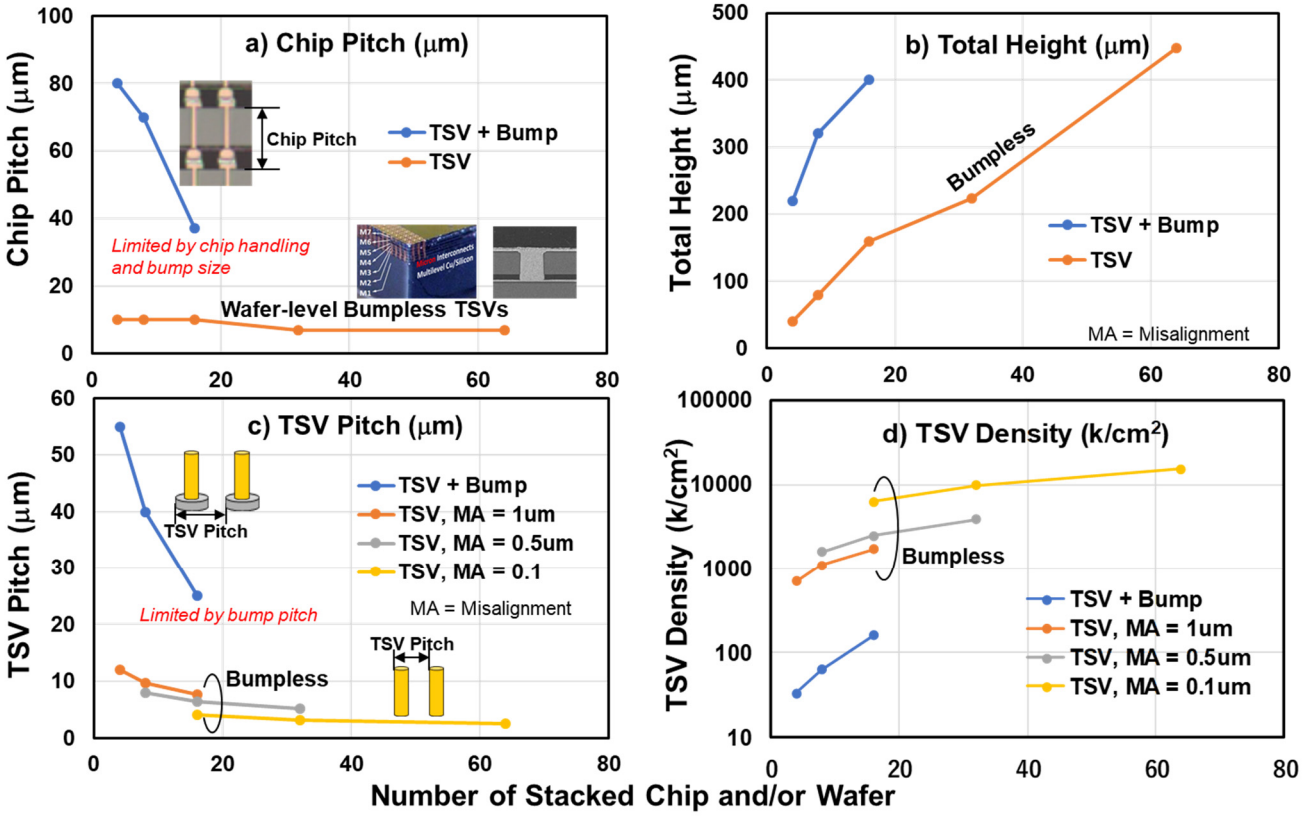

Fig. 3. Trends of TSV interconnects as a function of number of stacked layers: a) chip pitch, b) total height, c) TSV pitch, and d) TSV density. For bumpless process, TSV diameter and space between pads were 9 to 1 μm and 1 μm, varied with misalignment (MA), respectively.

wafers. The TSV density ranged from 10^6/cm^2 to 10^7/cm^2, which is 10- to 100-times larger than the case of TSVs and bump interconnects.

4. Bandwidth comparison

By using high-density bumpless TSV interconnects, a large number of parallel channels with low impedance can be used to implement high-bandwidth memories (HBMs). Because parallel data transfer realizes a high overall throughput, the bandwidth increases with the number of channels. For example, bandwidths on the order of Terabyte/s or greater can be achieved, as shown in Figure 4. At the same time, since the high-density TSVs act as thermal tunnels, temperature increases are mitigated in 3D systems [7].

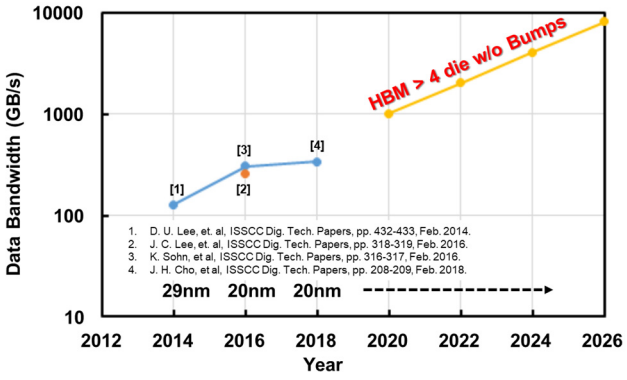

Fig. 4. Trend of data bandwidth of HBMs. Bandwidth increases with multi-stack having high-density parallel I/Os.

5. Conclusions

We have described a bumpless interconnect process and its prospects for realizing Terabyte bandwidth without sacrificing power efficiency, counter to conventional TSVs and bump interconnects. A combination of 2D scaling and 3D stacking, called co-integration, overcomes the problems associated with scaling, and a roadmap towards high-density integration backed up by production costs can be formulated.

Acknowledgements

The author would like to thank the WOW alliance members in Tokyo Tech.

References

[1] K. Sakui and T. Ohba, IEEE CICC Dig. Tech. 22-6, April 2019.
[2] K. Sakui and T. Ohba, IEEE IMW Dig. Tech., pp.60-63, May 2019.
[3] T. Ohba, Electrochem. Soc., Trans., **34** (1), pp. 1011-1016, 2011.
[4] T. Ohba, et. al, IEICE Electronics Express, **12** (7), pp. 1-14, 2015
[5] Y. S. Kim, et. al, IEEE VLSI Symp., pp. 22-23, 2014.
[6] Y. Mizushima, et. al, AMC/ADMETA 2017.
[7] H. Ryoson, K. Fujimoto, and T. Ohba, ICEP Dig. Tech. Papers, pp. 522-525, 2017.