Multilayer Stacking Technology Using Wafer-to-Wafer Stacked Method

NOBUAKI MIYAKAWA, EIRI HASHIMOTO, TAKANORI MAEBASHI, NATSUO NAKAMURA, YUTAKA SACHO, SHIGETO NAKAYAMA, and SHINJIRO TOYODA
Honda Research Institute Japan

We have developed a new three-dimensional stacking technology using the wafer-to-wafer stacked method. Electrical conductivity between each wafer is almost 100% and contact resistance is less than 0.7Ω between a through-silicon via (TSV) and a microbump. We have also created a prototype of a three-layer stacking device using our technology, where each wafer for the stacking is fabricated by using 0.18um CMOS technology based on 8-inch wafers. The device is operated by two times the frequency of the multichip module (MCM) device case using a two-dimensional device with identical functions and minimally different power consumption. The yields obtained from the results comprising all functional tests are over 60%.

Categories and Subject Descriptors: B.2.2 [**Arithmetic and Logic Structures**]: Performance Analysis and Design Aids—*Verification*; B.2.4 [**Arithmetic and Logic Structures**]: High-Speed Arithmetic—*Cost / Performance*; B.4.3 [**Input/Output and Data Communications**]: Interconnections—*Physical structures*; B.4.5 [**Input/Output and Data Communications**]: Reliability, Testing, and Fault Tolerance—*Hardware reliability*; B.7.1 [**Integrated Circuits**]: Types and Design Styles—*Advanced technologies*; B.8.0 [**Performance and Reliability**]: General

General Terms: Design, Performance, Reliability, Verification

Additional Key Words and Phrases: 3D integration, stacking process, design, hardware

ACM Reference Format:

Miyakawa, N., Hashimoto, E., Maebashi, T., Nakamura, N., Sacho, Y., Nakayama, S., and Toyoda, S. 2008. Multilayer stacking technology using wafer-to-wafer stacked method. ACM J. Emerg. Technol. Comput. Syst. 4, 4, Article 20 (October 2008), 15 pages. DOI = 10.1145/1412587.1412593 http://doi.acm.org/10.1145/1412587.1412593

1. INTRODUCTION

Many research organizations have been investigating stacking technology, both in the areas of assembly technology and device technology [Koyanagi et al. 1998;

Authors' addresses: N. Miyakawa (Director and Project Manager), E. Hashimoto (Senior Researcher), T. Maebashi (Researcher), N. Nakamura (Senior Researcher), Y. Sacho (Guest Engineer), S. Nakayama (Chief Engineer), S. Toyoda (Principal Researcher), Honda Research Institute Japan Co., Ltd., 8-1 Honcho, Wako-shi, Saitama 351-0188, Japan; email: {miyakawa, hashimoto, maebashi, natsuo_n, y-sacho, nakayama, s.toyoda}@jp.honda-ri.com.

SEMANTIC 2004]. For example, regarding assembly technology using stacking technology. Investigations have included board assembly such as MCM, package-on-package (PoP), and multichip package (MCP), as well as chip stacking assembly such as 3D system-in-package (SiP). In the case of stacking device technology, most researchers have been developing this technology using TSV (through-silicon vias) [Schaper 2004], and there are many trials underway for stacked process methods (chip-to-chip, chip-to-wafer, wafer-to-wafer). These process methods are proposed to explore trade-offs among different design goals. For example, in the case of chip-to-wafer stacking technology, a known good die (KGD) is connected to a good chip on another wafer [Morrow et al. 2006; Fukushima et al. 2005] such that the overall yield can be improved, at the expense of extra testing cost for KGDs. On the other hand, wafer-to-wafer stacking methodology is good for high-throughput manufacturing [Maebashi et al. 2007], with the disadvantage of lower yield compared to the chip-to-wafer stacking method.

3D stacking technology is very promising as a future technology because of its many benefits, such as its high performance, low power, smaller footprint, and enabling of heterogeneous technology integration [RTI 2004–2007; Iwata et al. 2005; Topol et al. 2005; ISSCC Forum 2007]. Such benefits of the 3D stacking technology will be useful for those computer systems in which there are multiple connection paths between each chip. For example, the interface circuits of the stacked device will be miniaturized, and can operate at high speed because of the shortered wire length for connections between layers. Wire length is also optimized by using the connection of the vertical direction in the case of bus-width expansion. Moreover, the propagation-delay time between each bit can be affected almost equivalently by using this stacking technology. And there is a potential for its application in the new technology environment of computer systems of the near future, by means of a fusion between stacking technology and post-CMOS technology, such as nanotechnology, biotechnology, and quantum devices.

Even though there is much research on 3D integration technology using TSVs such stacking device technology is not yet well on its way to industrial manufacture. The main reason is a lack of actual results with large-scale wafers, for example, the relationship between yield data and stacked layers, and also the actual data effected through using stacking technology. Moreover, to expand the application of the technology designers are posed with new challenges such as layer division, 3D floor planing, and clock tree synthesis for 3D systems.

In this article, we introduce a new 3D stacking technology using the wafer-to-wafer stacked method. We also demonstrate our technology using a prototype of a three-layer stacking chip, in which each wafer for the stacking is fabricated by using 0.18um CMOS technology based on 8-inch wafers. The device is operated by two times the frequency of the multichip module (MCM) device case, using a 2D device with identical functions and minimally different power consumption. The yields obtained from the results comprising all functional tests are over 60%.

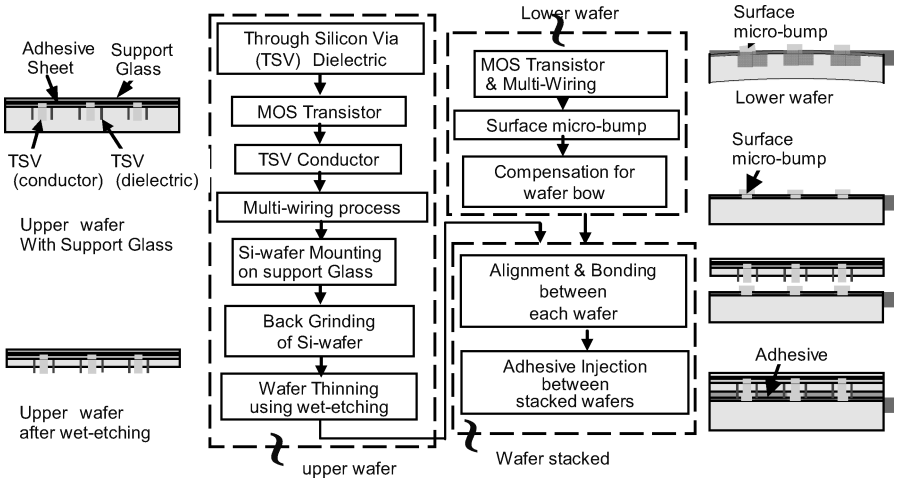


Fig. 1. Proposed stacking process.

2. PROPOSED STACKING PROCESS

We have investigated wafer-to-wafer stacking technology using large-scale wafers. Our stacking process comprises three steps, as shown in Figure 1, that is, upper-wafer fabrication, lower-wafer fabrication, and wafer stacking processes.

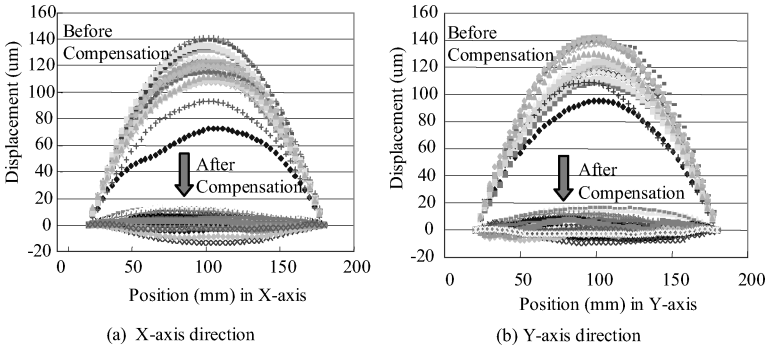
Our TSV processes have two steps before/after the MOS transistor fabrication process, because each fabrication process for the TSV dielectric and TSV conductor depends on the temperature. After multiwiring, the wafer is mounted on a support glass for wafer thinning.

The wafer thinning process has two steps: back grinding and wet etching. Note that this technology adds a wet-etching process to the conventional grinding and polishing process of an upper wafer. The cross-section of the upper wafer after the wet-etching process is also shown in Figure 1.

The lower-wafer fabrication process uses the same processes as for the upper wafer. However, the lower wafer has a bow due to multiwiring, as it is not mounted on a supporting glass after fabrication. This bow should be compensated for by the alignment of the stacking process. Figure 2 shows a comparison before and after this compensation. Therein it can be seen that, after some surface microbumps on the lower wafer are formed, our process remedies the wafer bow problem.

After preparation of the wafers for stacking, the wafer stacking process starts with alignment and bonding between each wafer, and then an adhesive is injected between stacked wafers.

Let us compare the connection portion of many previous cases and ours (see Figure 3): On the left side, the number of factors for total resistance is five and the resistance value dispersion between layer bumps increases for point-to-point connections. On the right is the cross-section and total resistance value of our connection method, reflecting a direct stacked method wherein the TSV



(a) X-axis direction (b) Y-axis direction

Fig. 2. Compensation for wafer bow.

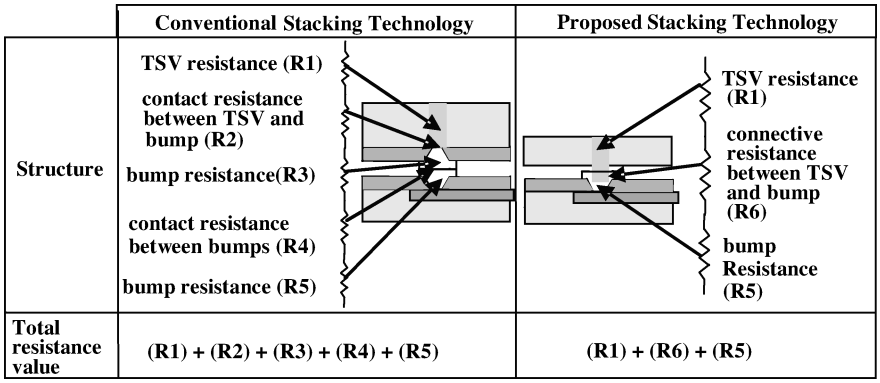
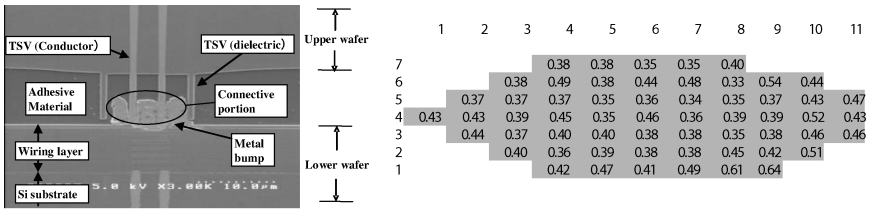


Fig. 3. Connection portion between TSV and microbump.

of an upper wafer is directly connected with a surface microbump of a lower wafer. Here, the number of factors for total resistance is three, while the resistance value dispersion between each layer is small and stable for face-to-face connection.

Our process does not need back-side microbumps, therefore the total process length is simplified and the reliability of the device is improved. Additionally, a unique TSV structure has been investigated for the wet-etching and stacking processes. The dielectric is fabricated by SiO2/poly-si and the conductor is fabricated by burying tungsten, after the trench hole for the TSV structure is opened by Bosch deep-trench-etching technology. (A Bosch process is a deep-trench-etching method with a high aspect ratio that alternates a deposition step for sidewall protection and an etching step. The angle of the sidewall is processed at almost 90 degrees to the surface and deep etching can be achieved.) In addition, the connection process is shortened.

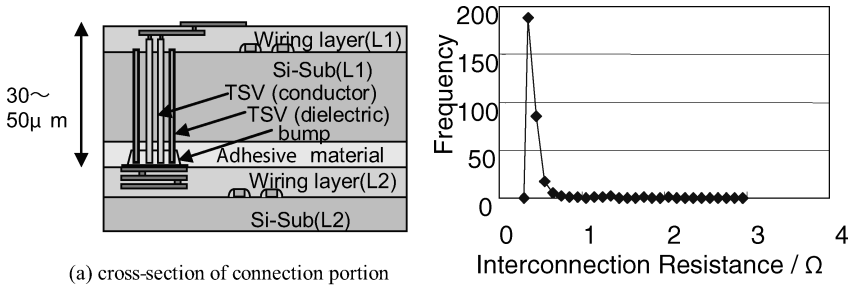
A photograph of the direct connected portion between the TSV of an upper wafer and the surface microbump of a lower wafer can be found in Figure 4(a). We evaluated the data of the electrical connectivity between TSV and microbump using TEG chips for an 8-inch wafer, which is derived from the whole-wafer interconnection resistance of a stand-alone case after two-layer stacking.



(a) cross-section of connection portion between TSV and microbump

(b) contact resistance between TSV and microbump unit: Ω (measured at TEG chip in same wafer)

Fig. 4. Cross-section of connection portion and connectivity resistance.



(a) cross-section of connection portion

(b) relationship between frequency distribution and resistance

Fig. 5. Frequency distribution of resistance with six stacking wafers.

As a result (shown in Figure 4(b)), the electrical connectivity between each layer approaches 100%, and almost all values of the stand-alone case are less than 0.7 ohm.

After tests with 6 stacked wafers for the frequency distribution of electrical conductivity, the connection-resistance dispersion is small and stable, with a very sharp frequency-distribution curve, as detailed in Figure 5.

Evaluation results of a 30- and 300-chain case, after two-layer respective stacking (as indicated in Figure 6), show that electrical connectivity is stable and frequency distribution is small. However, this data includes the parasitic resistance values for the connection between each chain.

We measured the electrical characteristics between TSV and microbumps using the 300-chain case, and the resistance value and capacitance of the TSV are detailed in Figure 7. The mean of resistance value is 1.5 ohms per chain by including the parasitic resistance value for the wiring, and the capacitance value of a TSV is less than 3pF per TSV.

The characteristics of our stacking method are as follows.

—Stacking technology is wafer-to-wafer.

—All back-side processes are removed, except for wafer thinning.

—Wet etching is used for wafer thinning.

—The total stacking process is simplified and shortened.

—TSV and microbumps are directly connected.

This technology is applicable for larger-scale wafers of more than 8 inches.

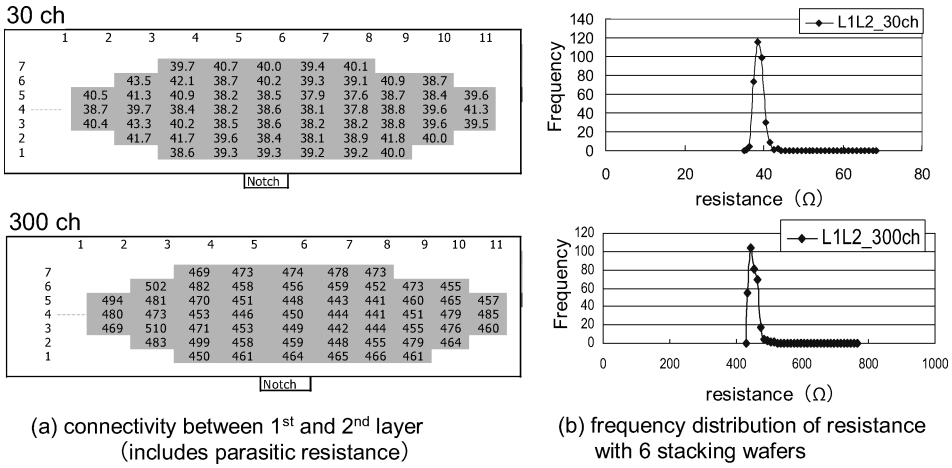


Fig. 6. Electrical connectivity between two layers.

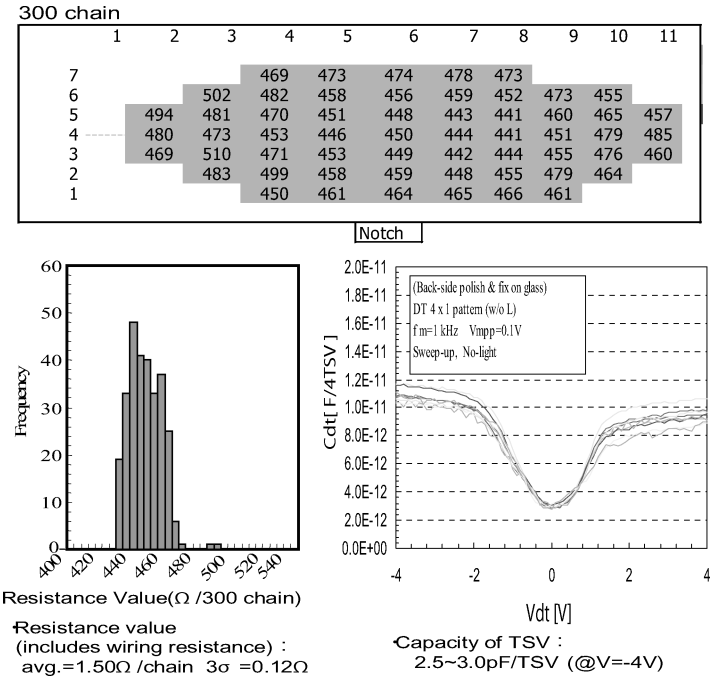
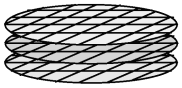


Fig. 7. Characteristics of TSV and microbump.

3. PROTOTYPE OF A THREE-LAYER STACKING DEVICE

We have tried the prototype of three-layer stacking devices using our technology. The trial manufacture has verified that our technology in fact incorporates the major distinguishing characteristics of 3D stacking. The 3-wafer stack of Figure 8 reveals that each wafer is structured by a microprocessor (L1); custom circuits, including analog circuits (L2); and a 64Mbit SDRAM (L3), respectively.



	Main function
1st Layer	Processor (SH4)
2nd layer	Custom Circuits included in AD Converter
3rd Layer	64Mbit SDRAM

Fig. 8. Structure of stacked devices.

Table I. Contents of Wafer-to-Wafer Stacking

Item	Contents
Wafer size	8 inches
Chip size	8.44×4.69 mm^2
Number of stacked layers	3 layers
Number of TSV between each chip	L1 –L2: 1,056 /chip L2 –L3: 287 /chip
Number of TSV between each wafer	L1 –L2: 655,584 /wafer L2 –L3: 180,160 /wafer
Process	0.18 μm CMOS

Also, the number of TSVs between L1- and L2 wafers is 655,584 and between L2- and L3 wafers is more than 180k (see Table I). It contains over 2Mgates and 64Mbit memory. The device size is 8.44mm $\times$ 4.69mm, with more than 1400 TSV wires. For greater detail, the photo-graphs of Figure 9 feature the face of each layer and the cross-section, respectively, of the die after stacking.

We designed the trial material according to this design flow as shown in Figure 10 after we prepared the 2D design data according to the target specifications. The first step was the transformation from 2D data to 3D data which distributes each layer. After this transformation we checked the equivalence of the logic before/after transformation. The next step was the distribution of circuits and clock signals in each layer and the extraction of the critical path, and after that the optimization of the critical path. We have checked the logic equivalence before/after layout.

Subsequent to this we did the timing analysis, and compared the original 2D design data with the converted data from 3D data of all layers to 2D data. Next, the 3D design data was modified and simulated until its performance and power consumption met the target specifications. The comparison results paralleled the target specifications in all of the modifications. Finally, we did the tape-out after the data successfully met the criteria.

The measurement structure of the three-layer stacking case from L1 to L3 and the data of the frequency distribution of the electrical conductivity results with six stacked wafers are shown in Figure 11. The frequency distribution is less than 6 ohms and the distribution curve is sharp: the same as with the two-layer-stacking case, though that data includes the parasitic resistance values for the connection.

Our connection method is very useful for the stacking technology and the electrical characteristics of the connection portion is remarkable, as noted from Figure 5 and Figure 11.

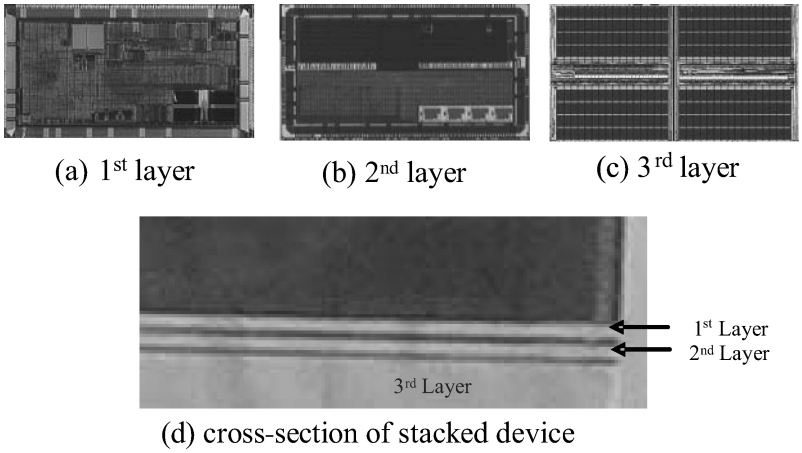


Fig. 9. Device photograph of layers1, 2, and 3, and cross-section photograph of stacked device.

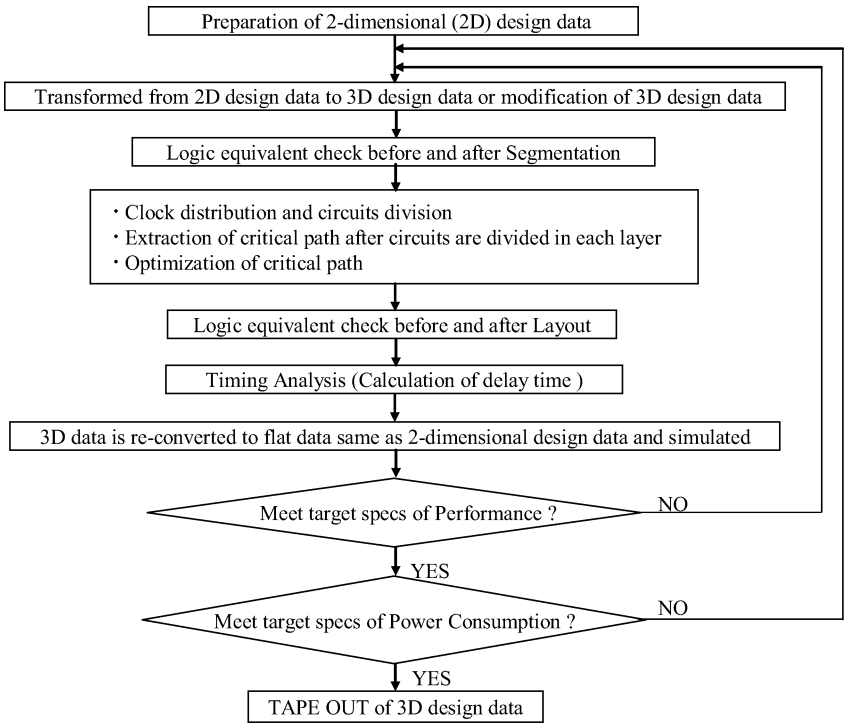
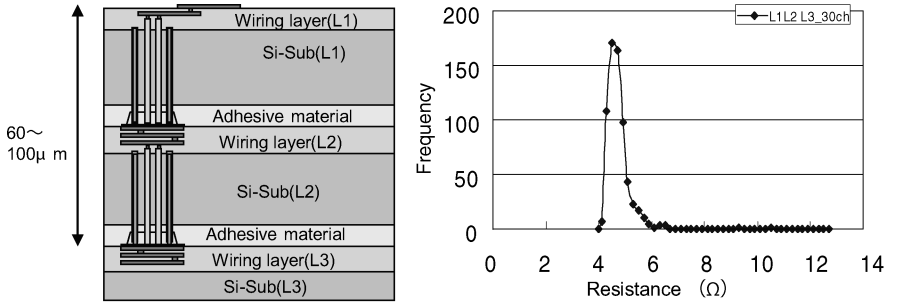


Fig. 10. Design flow of a 3D device.

We evaluated the basic characteristics by the probe test after 3-wafer stacking. Aqua chips show pass chips, pink chips show failed chips, and gray show TEG chips. The yields gotten from the results comprising all functional tests are more than 60%, as shown in Figure 12.



(a) Cross section of connection portion

(b) Relationship between Frequency distribution and resistance

Fig. 11. Frequency distribution of resistance with six stacking wafers.

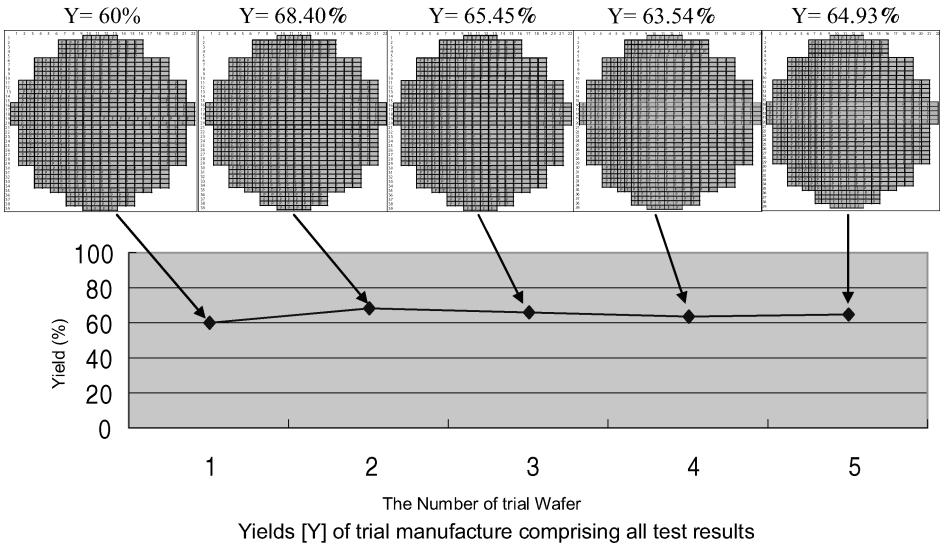


Fig. 12. Probe examination results of trial devices.

The electrical connectivity between each wafer is superior to former experiments; such values suggest a high probability for industrial fabrication in the near future, although some further investigation remains to be done.

We measured the electrical performance of the first layer, the so-called "shmoo test," using a tester after packaging. We got results such as those of Figure 13. This data includes test results of three kinds of devices, such as best case, typical case, and worst case. The second column is the supply voltage of the core and next column is the supply voltage of the I/O interface. The temperature for the test is 10° and 85°C. The operational frequency is from 278MHz to 217MHz at the typical supply voltage. This value is better than that of the MCM device case using 2D devices.

In our trial manufacture, pipelined 10-bit AD converters were placed in the L2 layer after the performance of AD converter was checked by using TEG. We

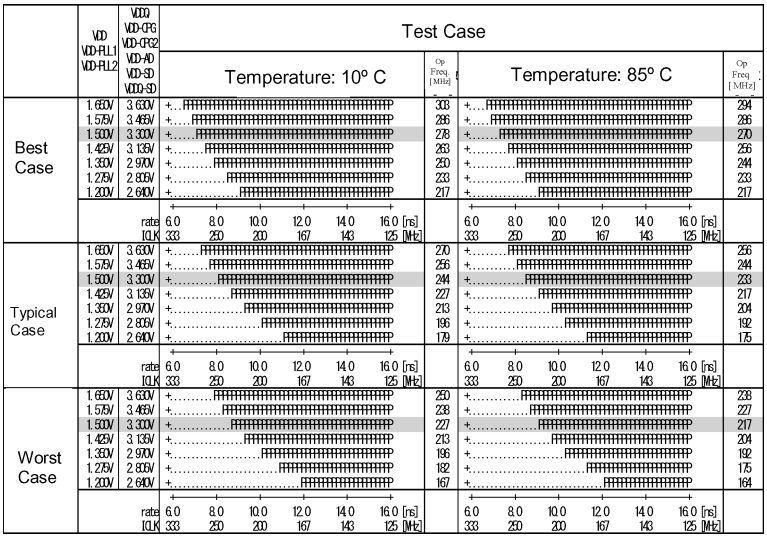


Fig. 13. Evaluation results of trial manufacture using tester.

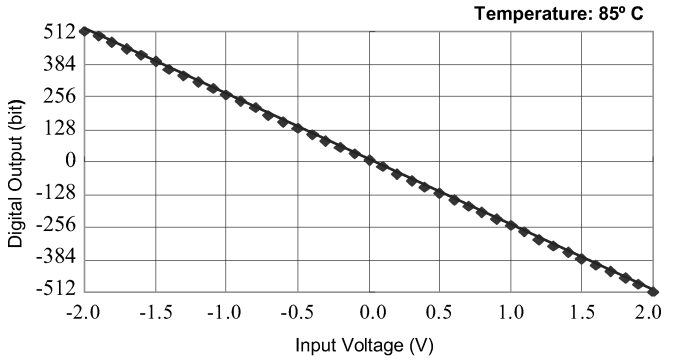


Fig. 14. Test results of 10-bit AD converter of second layer.

achieved good results (see Figure 14) without stress dependency of the adhesive material. Consequently, it may not be problematic to place any layer same as in analog circuits, although it depends on the characteristics of a differential amplifier which is influenced by the residual stress of the adjacent wafer, bonding pressure at the stack and thermal treatment at the adhesive injection.

We evaluated the relationship between operational frequency and power consumption using the trial manufacture and MCM device having the same function (see Figure 15). The red line shows the relationship of the trial manufacture. The blue line shows that of MCM device. In comparing both devices, the trial device has twice the operational frequency and less than one-third the power consumption of the MCM device under the same test conditions.

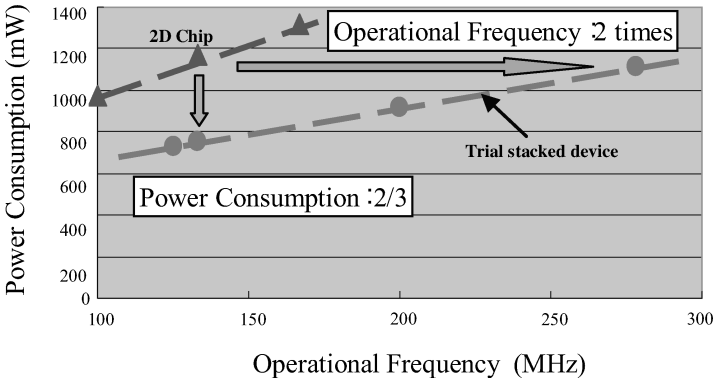


Fig. 15. Relationship between operational frequency and power consumption.

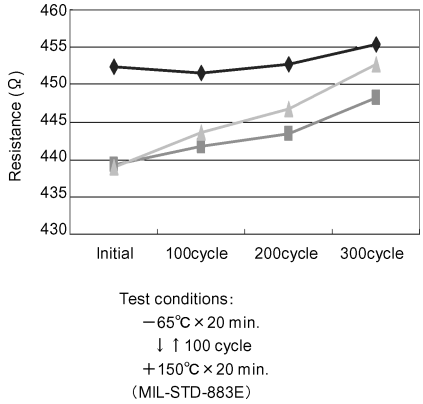


Fig. 16. Reliability test results (Mil-Spec standard).

The power consumption data may decrease more because the SDRAM wafer had circuits for an I/O interface.

We obtained the reliability test results of our device during the thermal cycle test. The results under the Mil-Spec and Consumer-Spec, standards are indicated in Figure 16 and Figure 17, respectively. It can readily be seen that our trial manufacture has no problem with structure from a reliability viewpoint.

4. ISSUES FOR COMPUTER SYSTEMS

There are some open issues for realizing computer systems that apply 3D stacking technology. First is the expansion method of the bandwidth for the memory: specifically, how to improve the placement density of TSV.

Second is the relationship between chip size and the number of interface terminals. The chip size of a device fabricated by using stacking technology becomes smaller. However, the chip must keep a given number of input and output terminals, respectively, in order to connect with the additional memory and peripheral circuits. We need new assembly technology for this.

A third challenge is the countermeasure for thermal awareness. Until now, it has been put forward in some technical papers that the heat generation of each

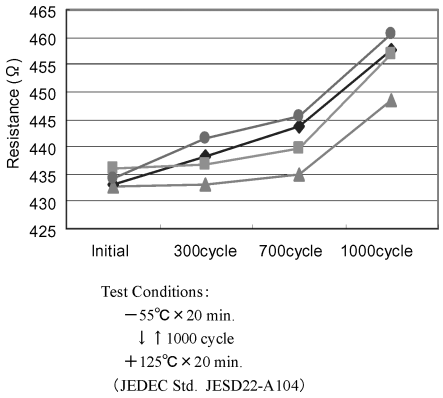


Fig. 17. Reliability test results (Consumer-Spec standard).

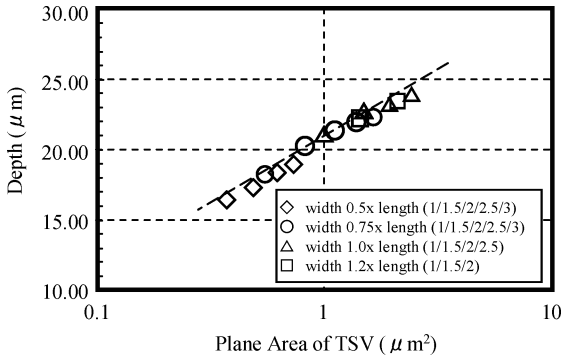


Fig. 18. Relationship between depth and plane area.

stacked chip is uniform [Xie et al. 2006]. While this may be the case, a more detailed analysis is needed here. Specifically, when the current is concentrated at a particular point, the temperature of this device will go up. Moreover, device mobility will decrease, thereby affecting performance, with this temperature increase. Another problem is the danger of fire, which, finally, can bring fatal damage to this 3D device. So we must investigate such "hot spot" phenomena and develop appropriate countermeasures.

We have investigated the etching of the small-size TSV and this investigation is useful for the extension of the bandwidth.

TSV is a connection wire between each chip and wafer in the 3D stacking technology; the thickness of the stacked chip and wafer is decided by means of the relationship between the size of the TSV and the depth after etching, and the resistance value is decided by the buried material. The etching of TSV is done by the Bosch method, which repeats the etching and the film deposition for the protection of the sidewall after patterning for taking the vertical direction. The etched shape is influenced by both ion-etching with the anisotropic characteristics and radical-etching with isotropic characteristics. The plural patterns of TSV are etched to get the relationship between a plane area and the depth. We show in Figures 18 and 19 each relationship. There is a logarithmic relation

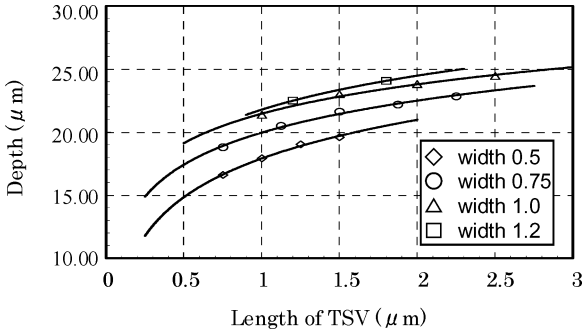


Fig. 19. Relationship between depth and TSV length.

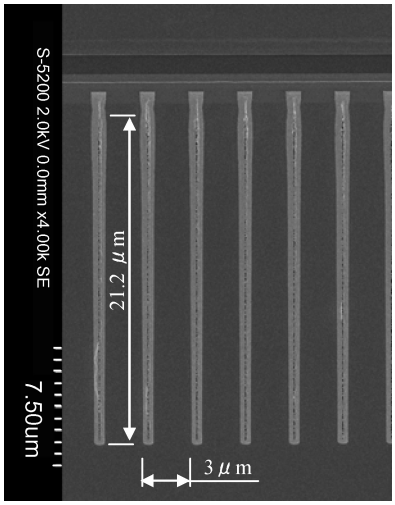


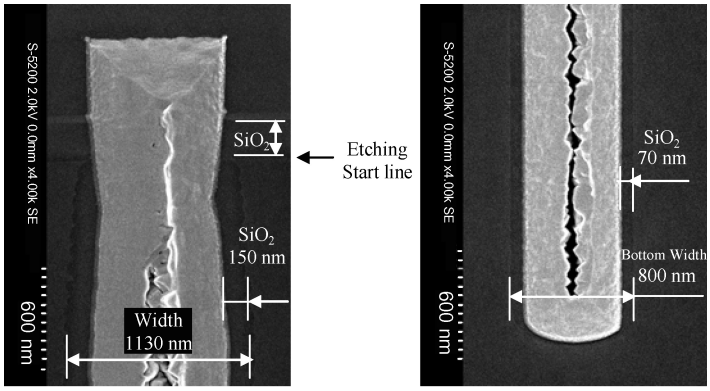
Fig. 20. 1 μmΦ TSV placed at 3 μm pitch.

between the plane area and the depth from both figures and this has nothing to do with the shape of the plane area.

We did the experiments with small plane area based on the results of Figures 18 and 19. The cross-section of TSV, where the tungsten is buried in the hole after etching a plane area of 1 μm^2 and placed between 3 μm pitch, is shown in Figure 20. It is not buried completely and creates a slit because the open portion of the top is closed. However, the stacking process is not influenced by keeping both the strength of TSV and the lower resistance value.

The etching depth of the size is over 21 μm, as shown in Figure 18, and the widest portion which is influenced by the reflection of etching ion near the top is more than 1100 nm, as shown in Figure 21(a). Furthermore, the width of the bottom portion is near 800 nm, as shown in Figure 21(b). The difference in width between the widest portion and the bottom is about 300 nm and the vertical angle of the TSV is kept over 89.5.

Computer systems need the larger bandwidth of the bus and a reduction the signal delay distortion among in bus signals. Therefore, TSVs have to be



(a) near the top portion (b) the bottom portion

Fig. 21. Top and bottom of 1 μmΦ TSV.

small and to have the placed pitch narrow. However, it is important that the characteristics of an active device like a MOS transistor are not influenced by the defects and stress induced by the etching process.

ACKNOWLEDGMENTS

The authors would like to thank Hitachi, Ltd. Micro Device Division for providing wafer fabrication using 0.18 um CMOS technology based on 8-inch wafers and for the evaluation of test wafers and test devices. The authors would also like to thank Soliton Systems K.K. for the stacked data conversion, Renesas Technology Corp. (formerly SuperH, Inc), Hitachi Information & Communication Engineering, Ltd., and Hitachi ULSI Systems Co. Ltd., for the design of SH4. The authors also would like to thank N. Horino for his support.

REFERENCES

FUKUSHIMA, T., YAMADA, Y., KIKUCHI, H., AND KOYANAGI, M. 2005. New three-dimensional integration technology using self-assembly technique. In *IEEE International Electronic Devices Meeting Conference Digest Technology Papers,* 359–362.

ISSCC. 2007. *ISSCC Forum on the Design of 3D-Chipstacks.*

IWATA, A., SASAKI, M., KIKKAWA, T., KAMEDA, S., ANDO, H., KIMOTO, D., ARIZONO, D., AND SUNAMI, H. 2005. A 3D integration scheme utilizing wireless interconnections for implementing hyper brains. ISSCC Dig. Tech. Papers, 262–263.

KOYANAGI, M., KURINO, H., LEE, K. W., SAKUMA, K., MIYAKAWA, N., AND ITANI, H. C. 1998. Future system-on-silicon LSI chips. *IEEE Micro. 18*, 4, 17–22.

MAEBASHI, T., NAKAMURA, N., NAKAYAMA, N., AND MIYAKAWA, N. 2007. New fabrication method for multi-layer stacked devices using wafer-to-wafer stacked technology based on 8-inch wafers. *ESSDERC*, 251–254.

MORROW, P. R., PARK, C.-M., RAMANATHAN, S., KOBRISKSY, M. J., AND HARMES, M. 2006. Three-Dimensional wafer stacking via Cu-Cu bonding integrated with 65-nm strained-Si/Low-k. *IEEE Electron. Devices 27*, 5 (May).

RTI. 2004–2007. *Proceedings of the 3D Architectures for Semiconductor Integration and Packaging Conference.*

SCHAPER, L. 2004. Copper electroplated through silicon vias for 3-DIC interconnection. In *Proceedings of Advanced Metallization Conference, MINI-Workshop*.

SEMANTIC. 2004. *Proceedings of the 3D Technology, Modeling, and Process Symposium*.

TOPOL, A. W., LA TULIPE, D. C., SHI, L., ALAM, S. M., FRANK, D. J., STEEN, S. E., VICHICONTI, J., POSILLICO, D., COBB, M., MEDD, S., PATEL, J., GOMA, S., DIMILIA, D., ROBSON, M. T., DUCH, E., FARINELLI, M, WANG, C., CONTI, R. A., CANAPERI, L., DELIGIANNI, A., KUMMAR, A., KWIETNIAK, K. T., D'EMIC, C., OTT, J., YOUNG, A. M., GUARINI, K. W., AND LEONG, M. 2005. Enabling SOI-based assembly technology for three-dimensional (3D) integrated circuits (ICs). In *IEEE International Electronic Devices Meeting Conference Digest Technology Papers,* 363–366.

XIE, Y., LOH, G., BLACK, B., AND BERNSTEIN, K. 2006. Design space exploration for 3D architectures. *ACM J. Emerg. Technol. Comput. Syst. 2*, 65–103.

Received June 2008; accepted July 2008