On-Chip TSV Testing for 3D IC before Bonding Using Sense Amplification

Po-Yuan Chen†‡ and Cheng-Wen Wu†
†Department of Electrical Engineering
National Tsing Hua University
Hsinchu, Taiwan 30013, ROC
Email: {pychen, cww}@larc.ee.nthu.edu.tw

Ding-Ming Kwai†
‡SoC Technology Center
Industrial Technology Research Institute
Hsinchu, Taiwan 31040, ROC
Email: dmkwai@itri.org.tw

Abstract - We present a novel testing scheme for TSVs in a 3D IC by performing on-chip TSV monitoring before bonding, using a sense amplification technique that is commonly seen on a DRAM. By virtue of the inherent capacitive characteristics, we can detect the faulty TSVs with little area overhead for the circuit under test.

Key-words: 3D IC, Defect-Based Testing, Interconnection Test, Sense Amplifier, TSV.

I. INTRODUCTION

Recent research and development in three-dimensional integrated circuit (3D IC) has been emphasizing the benefit of packing density which can be attained by stacking a growing number of dies. Additionally, 3D IC offers an opportunity to integrate heterogeneous processes in a more efficient manner, improve speed performance with smaller interconnect delays, decrease power consumption with shorter wire lengths, and increase data bandwidth by using short vertical links called through silicon vias (TSVs) [1–4]. Potential applications include processor (stacking CPU and various levels of caches) [5–6], memory (stacking SRAM, DRAM, and/or Flash) [7–9], wireless sensing [10], etc. Although there exist alternatives, like wire bonding and micro-bumping, using TSVs achieves higher interconnection density and better performance [11–14]. According to the step of TSV formation in an overall 3D IC manufacturing sequence, we could classify TSV technologies into two main categories, namely, via-first and via-last. One categorization is to separate by the bonding step: The via-first processes form the TSVs on each wafer prior to the bonding step, and the via-last processes form the TSVs after [4, 15].

In spite of many advantages, there are some problems associated with 3D IC. One of the most important issues is the compound yield loss due to die stacking [16]. To guarantee the stacking yield, the interconnections must be tested. The current interconnection test proposed for 3D IC is done with two or more dies in a stack, which is good only for TSVs after bonding [17–19]. Essentially, after two dies are bonded, the TSVs can be connected serially to form a daisy chain in an electric test or connected with flip-flops to form a scan chain in a structure test. Two possible configurations of scan-chain test are show in Fig. 1(a) and (b), respectively. In Fig. 1(a), each TSV is connected to a multiplexer and a flip-flop which can be used as a scan chain [18]. In Fig. 1(b), we reuse the flip-flops in the scan test and treat the TSVs as interconnections between wrappers, just as what we do for boundary scan in a planar design. There needs high reliability TSV channels, for test control or scan path. With the same test circuit in each layer, they can be tested in a complete or partial stack [19].

There are some limitations in these test schemes. First, they cannot be performed before bonding. A straightforward way for an electric test uses a daisy-chain structure of by alternate routes of TSVs on both the front and back sides of the wafer [20]. Apparently, this scheme is suitable only for the wafer acceptance test (WAT), since it is extremely difficult, if not impossible, to dismantle and rework the back metal once the TSV test is done. As a result, the observation of TSV failures at this stage relies solely on a couple of test keys on the scribe line. Second, individual TSVs are indistinguishable in a serial scan chain [17] or a daisy chain [20], so diagnosis becomes an issue. Probing both ends of a TSV can measure its resistance as the pass/fail criterion, but the area overhead for direct access is high, and thus is limited to a small number of sparse TSVs. Also, in general, for a die before bonding, the TSVs have one end on the backside that is not only floating but also buried deeply in the wafer substrate before thinning. Third, in the case of a via-first process, which intend to provide an interconnection density as high as $10^7$mm$^{-2}$, on-chip TSV monitoring becomes necessary. However, there are not always flip-flops connected to both ends of each TSV. Finally, the TSV failure rate affects the final yield exponentially with the number of dies in a stack. Unfortunately, it remains relatively high (>10 ppm) [18]. Without screening out the bad ones, the overall yield of the die stack will be low.

Fig. 1. Testing TSVs by reusing scan flip-flops (a) in serial connections or (b) in parallel connections.
What can we do to overcome the limitations mentioned above? Since the trend of scaling seems unavoidable and the yield of TSVs has a dramatic influence, we need some testing scheme which is performed before bonding in order to reduce the risk of bonding dies having irreparable TSV failures. Such a scheme also needs to distinguish fault-free and faulty TSVs efficiently so that the redundancy or repair strategy can be implemented with acceptable overhead. In this paper, we propose a method utilizing the parasitic capacitance of the TSV as an alternative of resistance monitoring that requires double-end probing. We focus on the via-first TSVs, because they are buried under metal layers and have higher density than the via-last ones. They cannot be tested by a scan chain, especially before wafer thinning and back metal patterning. Utilizing the capacitive characteristics of the TSVs, we can detect the faulty TSVs with little area overhead from the testing circuit. On-chip and parallel tests can be achieved, too.

The organization of this paper is as follows. In Section II, we briefly introduce the via-first TSVs and the associated defect types in the conductor and insulator, respectively. In Section III, we propose a new method to test the TSVs before bonding and present the flow of a discharging test. With such a test strategy, single-layer TSV monitoring can be performed on-chip using sense amplification. In Section IV, we illustrate the block diagram and schematic of the test circuit to show that it can be realized with very low area overhead. Simulation results and limitations of this scheme are discussed. In Section V, conclusions and future work will be given.

II. VIA-FIRST TSV AND DEFECT TYPES

A. Via-First Formation of TSV

We focus our targeted process on the via-first TSVs whose formations are performed before bonding. Fig. 2 depicts the cross sectional of a single via-first TSV before wafer grinding/thinning on the backside. It is most likely to be fabricated in a foundry and before the backend of line (BEOL) process for metal layers. One end of the TSV is connected to a metal layer. This layer can be any level varying from metal-1 to the topmost, based on different types of processes which all fall into the via-first category. The other end is floating and insulated from the substrate by a surrounding dielectric layer, as shown in the figure.

This kind of TSV could have a diameter smaller than 5μm, in order to achieve an interconnection density higher than 10^7/mm^2. Because the copper (Cu) process is widely used in silicon foundries, we assume in the following that the TSV is filled with Cu with barrier of titanium nitride (TiN), and the surrounding dielectric is silicon dioxide (SiO2). However, it is noted that the scheme proposed here can easily apply to other filling processes with different conductive materials, such as tungsten, nickel, and poly-silicon [4, 15].

B. TSV Equivalent Circuit

Considering a cylindrical TSV as shown in Fig. 2, we derive a lumped equivalent circuit with the following equations [21–22]:

$$R = \frac{R_{\text{Cu}}}{\pi d} \parallel R_{\text{bar}} = \frac{4\rho h}{\pi d^2} \| \frac{\rho_{\text{bar}} h}{\pi (d + t_{\text{bar}}) l_{\text{bar}}} \approx \frac{4\rho h}{\pi d^2}$$

$$C = \frac{2\pi h}{\ln((d + 2t_{\text{ox}})/d)} + \frac{\pi d^2}{4\varepsilon_{\text{ox}}}$$

where $h$ and $d$ are, respectively, the height and diameter of the Cu TSV, $t_{\text{bar}}$ is the thickness of the TiN barrier, and $t_{\text{ox}}$ is the thickness of the SiO2 insulator. The Cu resistivity is denoted as $\rho_{\text{Cu}}$, and the TiN resistivity as $\rho_{\text{bar}}$. The dielectric constant of SiO2 is $\varepsilon$. The value of $t_{\text{ox}}$ is about 60nm [23]. In other words, $d$ is much greater then $t_{\text{ox}}$. The lumped TSV resistance $R$ is hence the parallel combination of Cu resistance $R_{\text{Cu}}$ and TiN resistance $R_{\text{bar}}$. The contribution of $R_{\text{bar}}$ is negligible and can be ignored. This results in (1). Also, like a MOS capacitor, $C$ represents the maximum capacitance of the TSV under a low-frequency and high-voltage operation.

In (2), the first term on the right-hand side models the sidewall, and the second term the bottom disc. Note that if the wafer is thinned so that the vertical via can punch through the wafer, the back metal is deposited. In that case, the second term on the right-hand side of (2) is removed, but the parasitic capacitance is changed only slightly. To further simplify the modeling process, we assume that the pitch between two TSVs is large enough, so their coupling can be neglected.

C. Conductor Defects

We now examine some defect types that can occur in the TSV. Fig. 3 depicts a break in the conductor that induces an open failure. The break is located at a distance $h^*$ from the top end of the TSV. Note that the break may not be totally open, but instead it may behave like a resistive path. The defect, therefore, may be latent and requires burn-in or overstress to accelerate its failure mechanism. It can also occur between a contact and its corresponding TSV, which use different material.

With such a failure, the signal does not pass through from one end of the TSV to the other in specified time. The effective capacitance measured from the top end is reduced, which depends on how far the break is located. Clearly, the higher the break is located, the smaller the effective capacitance will be.
D. Insulator Defects

The insulator non-uniformity can be viewed as a result of non-conformal dielectric deposition or poor step coverage. This leads to a low breakdown voltage and a possible short between the TSV and the substrate. An example is shown in Fig. 4.

It may also be caused by particles induced by impurities or dust during the fabrication process. Such a defect occurs in the insulating layer and forms a bridge between the TSV and the substrate, as shown in Fig. 5. This is modeled as a resistive short defect. Such a resistive short defect creates a bypass to substrate and delays the signal, leading to slow transition time (a transition fault) or degraded substrate and delays the signal, leading to slow transition time (a transition fault) or degraded voltage level (a stuck-at-0 fault).

III. PROPOSED TSV TESTING METHODOLOGY

The concern now is how to test a single-ended TSV, i.e., before die/wafer stacking, when we are unable to test it as an interconnection just like the case after stacking. As mentioned in Sec. II, the open and short defects change not only the resistance but also the capacitance of a TSV. Hence, instead of trying to access both ends of the TSV, we measure its time constant (the product of \( R \) and \( C \)) variation to check whether it is fault-free or not. By this way, we treat each TSV as a DRAM cell and test it using sense amplification. Since DRAM sensing is inherently parallel, the scheme is capable of handling a larger number of TSVs.

The test flow of the proposed scheme is shown in Fig. 6. There are two phases to discharge the TSV after it is charged to \( V_{dd} \), with durations \( T_i \) and \( T_{ih} \), respectively. Assume an acceptable capacitance range of a TSV is [\( C_L, C_H \)], i.e., \( C_L \leq C \leq C_H \). Step 1 is the reset step, which executes a write-1 operation to all TSVs. Step 2 is the first discharge phase, a read-1 operation to check the lower bound \( C_L \). Note that TSV sensing does not require write-back as in DRAM. The read-0 operation in Step 4 follows a short hold-on (Step 3). Step 4 is the second phase discharge, which is used to check the upper bound \( C_H \). Again, the test sequence is \{write-1, read-1, hold-on, read-0\}. Step 3 is optional. It is executed when we need this time period to change the reference voltage of the sense amplifier to perform the read-0 operation in Step 4.

For example, consider a TSV with dimensions \( 5\mu m \times 50\mu m \), a dielectric thickness \( t_{\text{ox}} = 0.5\mu m \), and a nominal capacitance value of about 60fF. If we choose a variation range to be \( \pm 10\% \), then \( C_L \) is 54fF and \( C_H \) is 66fF. The equivalent circuit of the TSV and its discharge path are shown in Fig. 7(a). The \( RC \) model of the TSV is obtained by (1) and (2). The discharge path is modeled as a parasitic capacitance \( C_P \) and a channel resistance \( r_{on} \) of the MOS transistors which are cascaded with the TSV. In the figure, \( V_r \) is the input voltage of the sense amplifier. A discharge current \( I_{dis} \) is conducted from \( C \) through \( r_{on} \) to ground during the discharge phases. In the read-1 operation, by setting the sense amp enable signal \( \text{SAEN} \) for a period of \( T_i \), the pull-down circuit is turned on and \( I_{dis} \) pulls down the voltage level of the TSV. To de-
crease the parasitic effects, we need to place this test circuit as close to the TSV as possible.

From Fig. 7(b), we see that $C_L$ is the minimum capacitance that provides a voltage greater than the threshold $V_{th,H}$ after a period $T_L$ of discharge time; i.e., it passes the read-1 sensing test with $V_{th,H}$. The following read-0 senses the voltage again after a period $T_H$ of discharge time to check whether it is lower than the threshold $V_{th,L}$. The maximum capacitance that passes the read-0 sensing test with $V_{th,L}$ is $C_H$. In Fig. 7(b), the resolution of these discharge curves can be controlled by the thresholds of sensing voltages $V_{th,L}$ and $V_{th,H}$, the periods of discharge $T_L$ and $T_H$, and the on-resistance $r_{on}$ of the pull-down circuit.

Fig. 7. (a) Equivalent circuit of the proposed test scheme. (b) Threshold voltage of the TSV sensing and discharge time

Apparently, one limitation of this scheme is that some types of defects may not change the capacitance or resistance of the TSV. Consequently, no matter how fine the resolution is, we just cannot detect such faulty TSV by sensing. For example, a void (hollow hole) in the middle of a TSV, which is formed during Cu filling, may not be sensed if the resistance does not increase by a noticeable amount. This can lead to a reliability problem because of electromigration or stress migration. Although it is not detected here, it should be detected after certain stress tests and/or burn-in. Also, the variability of the capacitances and thresholds will affect the results above and should be considered when choosing the values of $C_L$ and $C_H$. If the capacitance range between $C_L$ and $C_H$ is too narrow, smaller discharge current or shorter discharge period is needed.

IV. EXPERIMENTAL RESULT

The block diagram of the testing circuit is shown in Fig. 8. Each TSV is connected to a test module (TM) and the normal function logic through a multiplexer (MUX). During the test mode, the MUX is switched to the test circuit and each TSV is controlled by its corresponding TM. First, the external tester activates the controller and input commands. The controller then takes control of the test procedure and generates the required signals to be broadcast to each TM. The sensing results are captured locally by flip-flops (FFs). TSVs can be tested in parallel all at once with a test circuit dedicated to each one, as shown on the left of Fig. 8. They can also be partitioned into disjoint groups, each shared with a test circuit. A TM is multiplexed by many TSVs, as shown on the right of Fig. 8. Even in the same group, interleaving testing is performed to minimize the crosstalk interference between adjacent TSVs. The outcomes are sent back to the test result collector (TRC). If the TSV failure rate is very high or very low, we can implement a compressor in the TRC to reduce the volume of test data. Since the number of test inputs and outputs is small, the testing procedure can be easily done by probing test pads during WAT before wafer thinning.

A more detailed schematic of the testing circuit based on the scheme presented above is depicted in Fig. 9. In order to minimize the area overhead, a tri-state buffer is used as the write driver, an NMOS transistor as the pull-down circuit, and two cascaded inverters as the sense amplifier, respectively. As discussed earlier, the write enable signal $WE$ is set in the write-1 operation, and the sense amp signal $SAEN$ is pulled up to turn the pull-down NMOS transistor on during the discharge period. Note that although there are two thresholds $V_{th,L}$ and $V_{th,H}$ in Fig. 7(b), the two inverters used here for sense amplification merge them into one threshold $V_{th}$ whose value is about 780 mV in our simulation, using the TSMC 0.18μm generic CMOS process.

Fig. 9. Schematic of a TSV testing circuit with two inverters performing the sense amplification.
The advantages of using inverters as the sense amplifier is its low area overhead, and that it can be merged into the subsequent logic circuit. The drawback is that the sensing margin can be low. As show in Fig. 10, there is only one threshold voltage $V_{th}$. Since the capacitance variations are small, the period $T_H$ is too small to be generated. The discharge capability is decided by the $W/L$ ratio of the pull-down NMOS transistor. We select the $W/L$ ratio to be 0.22μm/0.3μm, and the discharge periods $T_L$ and $T_H$ to be 0.5ns and 0.2ns, respectively.

![Fig. 10. Threshold voltage of the TSV sensing and discharge time with two inverters as the sense amplifier.](image)

In summary, the sensitivity and resolution of this scheme are affected by the discharge current, discharge time, threshold voltage of sense amplifier, and parasitic capacitance of the wires connected to the test circuit. If the discharge current is large, the voltage of the TSV may drop so quickly that it cannot be sensed correctly in time.

In our previous example with a 5μm × 50μm Cu TSV, the capacitance of the TSV is about 60fF and the variation range is ±10%. Simulation results are shown using waveforms in Fig. 11. There are five TSVs with capacitance varying from 48fF to 72fF in a step of 6fF. In the write-1 operation, all TSVs are set to a high voltage level. Then the read-1 operation discharges all TSVs for 0.5ns. After holding for 1ns, the final read-0 operation discharges all TSVs for 0.2ns. In the figure, the 48fF TSV fails the read-1 and the 72fF TSV fails the read-0. As mentioned earlier, the discharge time period $T_H$ is so short that it may not be easy for a simple circuit to generate the control signal. This is because the difference between the maximum and minimum capacitances is too small. One way to fix it is to change the hold-on operation to a write-1, and the second discharge period will then be increased to $T_H + T_L$.

We conduct an experiment so that we can take a closer look at the relationship between capacitance and discharge time period for different channel lengths of the pull-down NMOS transistor. In this experiment, the channel width is fixed at 0.22μm, which is the minimum in the 0.18μm process, and the channel length varies from 0.22μm to 0.55μm. The capacitances in these experiments are from 10fF to 200fF. These are the capacitances of TSVs, according to (2), with diameter varying from 1μm to 10μm and height from 30μm to 100μm. The thickness of the insulator remains at 0.5μm.

![Fig. 12. Discharge period versus maximum capacitance by sizing the pull-down NMOS transistor.](image)

Fig. 12 shows the maximum capacitance which can be discharged to ground within the period of time. Assume, e.g., the maximum capacitance is 120fF which can be discharged within 1.2ns, when the pull-down NMOS transistor has a channel length of 0.3μm. In this figure, if we choose the lower and upper bounds, $C_L$ and $C_H$, according to the size of the pull-down NMOS transistor, we can decide the discharge periods $T_L$ and $T_H$ in the read-1 and read-0 operations, respectively. By increasing the channel length to, say 0.55μm, we decrease the discharge current and increase the discharge period to 1.9ns, which eases the generation of the enable signal for the pull-down circuit.

V. CONCLUSIONS

We have presented a novel testing scheme for TSVs before bonding by sense amplification. The anomaly is detected by the aberration in charge or discharge time constant. With this scheme, the TSVs can be tested at the wafer level, and the identification of faulty TSVs before bonding becomes feasible. The defect types of a via-first TSV and its equivalent RC
model have been discussed. The test flow and implementation of the test circuit are very simple. If the ranges of resistance and capacitance of the TSV under test are changed, one can adjust the discharge current, threshold voltage of the sense amplifier, or the discharge time. With this scheme, some defective TSVs can be detected early before bonding, thus save the cost of the following futile actions. However, the process variations may change the RC, which will influence the offset voltage of the sense amplification circuit and the accuracy of the test. The coupling effect between TSVs and interconnects may also interfere the test. More detailed and thorough researches need to be done. In the near future, we will refine the design of sense amplification to increase its sensitivity and take the process variations and mismatches of transistors into consideration.

ACKNOWLEDGMENT

The results presented in this work would not have been possible without the helpful comments and suggestions from Y.-F. Chou and D.-S. Chen of the SOC Technology Center, Industrial Technology Research Institute, Hsinchu, Taiwan.

REFERENCES


