November 26, 2007
RT0771
Engineering Technology; Materials Science 9 pages

Research Report

Investigation of the thermal resistance of three-dimensional (3D) chip stack from the thermal resistance measurement and simulation of a single-stacked-chip

Keiji Matsumoto*, Kuniaki Sueoka, Katsuyuki Sakuma, Fen Chen (IBM S&TG) and Fumiaki Yamada

IBM Research, Tokyo Research Laboratory
IBM Japan, Ltd.
1623-14 Shimotsuruma, Yamato
Kanagawa 242-8502, Japan

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

Limited Distribution Notice
This report has been submitted for publication outside of IBM and will be probably copyrighted if accepted. It has been issued as a Research Report for early dissemination of its contents. In view of the expected transfer of copyright to an outside publisher, its distribution outside IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or copies of the article legally obtained (for example, by payment of royalities).

Investigation of the thermal resistance of three-dimensional (3D) chip stack from the thermal resistance measurement and simulation of a single-stacked-chip

Keiji Matsumoto*[1], Kuniaki Sueoka[1], Katsuyuki Sakuma[1], Fen Chen[2] and Fumiaki Yamada[1]
[1] : IBM Tokyo Research Laboratory, 1623-14 Shimotsuruma, Yamato-shi, Kanagawa-ken 242-8502, Japan.
[2]: IBM Microelectronics Division, Essex Junction, VT 05452
* Corresponding author. Phone : (+81)-46-215-5188. Fax : (+81)-46-273-7413. e-mail : keim@jp.ibm.com

Abstract

The thermal resistance of three dimensional (3D) chip stack is investigated by thermal resistance measurement and simulation of each component of a single-stacked-chip.

When multi-chips are stacked, the thermal resistance of each component of a single-stacked-chip adds up and influences the thermal resistance of 3D chip stack significantly. In this regard, the precise thermal resistance measurement and simulation of each component of a single-stacked-chip is important to understand the thermal resistance of 3D chip stack. The difficulty of measuring the thermal resistance of an interconnection by the laser-flash method has been shown by Yamaji et al[12]. In this study, steady-state thermal resistance measurement method employing liquid metal as contact material (which is applied between a sample and the measurement tool) is achieved and the thermal resistance of an interconnection is measured to clarify whether it is the thermal resistance bottleneck of 3D chip stack. Our measurements indicate the thermal resistance of a 200μm pitch, 9μm thick copper-tin (CuSn) interconnection[15] to be 0.078 ± 0.018 C cm^2/W. Also the thermal resistance of back-end-of-the-line (BEOL) of the 45nm technology node is simulated (finite-element method), by constructing an actual structural model and by assigning the experimental thermal conductivities to interlayer dielectrics (ILDs) of the model, to be 0.003 C cm^2/W. Further, the thermal resistance of a silicon substrate of 3D chip stack with various interconnection pitches is simulated, considering the concentrated heat flow to an interconnection, and then the correlation between the thermal resistance of a silicon substrate and the interconnection pitch is presented.

Based on the thermal resistance of the interconnection, the 45nm BEOL, the silicon substrate which are derived above, the total thermal resistance of 3D chip stack with the configuration of four-stacked chips is estimated. The dependence of the estimated total thermal resistance of the 3D chip stack on the interconnection pitch is presented. Also, combined with the thermal resistance of a cooling method (silicon microchannel cooler[23]), the maximum allowable power density of the 3D chip stack at the bottom of it is estimated and the dependence of the estimated maximum allowable power density on the interconnection pitch is described. The estimated maximum allowable power density of the 3D chip stack is compared with the ITRS prediction (cost-performance MPU maximum power density).

Keywords

Three-dimensional (3D) chip stack, thermal resistance, an interconnection, back-end-of-the-line, allowable power density, dependence on the interconnection pitch.

1. Introduction

As device scaling becomes more challenging, three-dimensional(3D) integrated-circuits(ICs) get more attention to enhance system performance without device scaling, owing to their higher interconnect density and shorter interconnect length. 3D ICs have attractive advantages as mentioned above, on the other hand, thermal performance of 3D ICs is a concern because of limited contact area with a cooling method.

Extensive simulation studies have been devoted to clarify the thermal characteristics of 3D chip stack in this decade. Chiang et al.[1] calculated the power consumption of 3D ICs by incorporating active layers and interconnect joule heating, and they analyzed the temperature rise of four different 3D logic-memory integration schemes. Similarly, Im et al.[2] analyzed the temperature rise of 3D ICs, in two bonding cases (The first one is wafer boding by a glue layer or solid-phase crystallization (SPC). The second one is wafer bonding by Cu pad thermo-compression) and in the second case they also investigated the dependence of the temperature rise of 3D ICs on the Cu pad area ratio. More recently, Puttaswamy et al.[3] calculated the temperatures of 2-layer and 4-layer stacked chips (the Alpha21364 processor) and they also assessed a thermal management solution of architectural level optimization[4]. Das et al.[5] calculated the temperature of stacked FFT(Fast Fourier Transform) datapath, depending on the number of stacks. In more detail, Hua et al.[6] performed thermal simulation of 3D ICs by considering the temperature dependence of transistor-delay and the temperature dependence of leakage power, including the effect of thermal-vias. The effect of thermal-vias were also approached by Chiang et al.[7-9] and Gplen et al. [10] and Yu et al.[11].

On the other hand, experimental studies on the thermal characteristics of 3D chip stack are very limited. Yamaji et al.[12] measured the contact thermal resistance between underfill and a chip by the laser flash method and based on the measurement results, they performed a numerical analysis of the thermal resistance of 3D chip stack. However, the difficulty of measuring the thermal resistance of an interconnection by the laser-flash method was shown and they pointed out that the careful attention had to be paid to uniform temperature distribution in specimen when applying the laser-

flash method to heterogeneous specimen, such as stacked chips with an interconnection. Sweet et al.[13] measured the thermal resistance of 3D multi-chip module by embedding polysilicon heaters and diode thermometers. They also performed the simulation using finite element method and compared the measured results with the simulated results. However, the difference between the measured results and the simulated results was attributed to the thermal resistance of the interconnection and it was not clarified.

As described above, although simulation studies on the thermal resistance of 3D ICs have been extensively performed, experimental results of the thermal resistance of 3D ICs are significantly lacking. As Gurrum et al[14] pointed out that an interconnection between a chip and a board is a key for the thermal characteristics of conventional 2D chip-to-board configurations, also in 3D chip stack, an interconnection between stacked chips is considered to be important for the thermal characteristics of 3D chip stack..

This study aims, firstly, to clarify the thermal resistance of an interconnection experimentally. Secondly, the thermal resistance of a silicon substrate, depending on the interconnection pitch, and also the thermal resistance of back-end-of-the-line (BEOL) is clarified by simulation. Based on the thermal resistance of each component (an interconnection, a silicon substrate and BEOL) of a single-stacked-chip, which is obtained above, the total thermal resistance of 3D chip stack is estimated. Further, combined with a presently available cooling method(silicon microchannel cooler[23]), it is estimated how much heat dissipation of 3D chip stack can be managed.

Again, in this study, the total thermal resistance of 3D chip stack is assumed to be composed of that of a silicon substrate, an interconnection, and back-end-of-the-line (BEOL). The thermal resistance of front-end-of-the-line (FEOL) is not considered here. Furthermore, with regard to silicon through via (TSV), as the structure of TSV is not yet determinate, therefore TSV is not taken into consideration.

2. Measurement method

When multi-chips are stacked, the thermal resistance of each component of a single-stacked-chip adds up and influences the thermal resistance of 3D chip stack significantly. In this regard, the precise thermal resistance measurement of each component of a single-stacked-chip is important to understand the thermal resistance of 3D chip stack.

Steady-state thermal resistance measurement method employing liquid metal as contact material (which is applied between a sample and the measurement tool) is achieved and is depicted in Figure 1. It consists of a heater, a cooler, two thermal sensors (T_1, T_2), and a sample which is sandwiched by two copper blocks with thermal sensors in them. Generated heat at a heater passes through a copper block and uniform heat distribution in horizontal direction is realized. Measurements are done in the conditions where heat flow to the outside of the measurement tool is decreased as much as possible.

Liquid metal (mixture of indium (In) and gallium (Ga), whose composition is indium of 75.5 % by mass and gallium of 24.5 % by mass) is applied between a sample and two copper blocks to realize a small contact thermal resistance between a sample and two copper blocks, and the resultant small fluctuation of the contact thermal resistance enables the precise measurement. In more detail, grease is usually used as contact material and its thermal conductivity is at best around 5W/mK, and it is hard to be thin because of its viscous property. Therefore, the contact thermal resistance of grease is as large as 0.19 C cm^2/W when it is measured in the method of Figure 1. and it fluctuates +-0.01 C cm^2/W depending on each measurement. This fluctuation prohibits from measuring the thermal resistance of a sample precisely. On the other hand, liquid metal (InGa) has high thermal conductivity (The thermal conductivity of indium is 82W/mK and that of gallium is 41W/mK) and is literally liquid at room temperatures, therefore it can be thin by being compressed. Moreover, it usually realizes an intimate contact with various surfaces. When an intimate contact cannot be obtained, very thin (in the order of 10nm) chrome and gold layer on surfaces improve the contact and the thermal resistance of these thin chrome and gold layer is negligibly small (in the order of 10^{-6} C cm^2/W) compared with that of a sample. These properties of InGa liquid metal lead to its small contact thermal resistance such as 0.008 C cm^2/W when it is measured in the method of Figure 1. and its resultant fluctuation is as small as +-0.001 C cm^2/W. The comparison of the contact thermal resistance and its resultant fluctuation between grease and liquid metal, which are measured by the method of Figure 1, is described in Table 1. When the reaction between copper and liquid metal is concerned, some barrier metals (e.g., nickel) should be applied on the surface of copper blocks with about 1μm thickness, and the thermal resistance of this barrier metal is calculated to be one order lower (in the order of 10^{-4} C cm^2/W) than that of liquid metal (0.008 C cm^2/W).

The thermal resistance of a sample is derived by Equation 1. In order to evaluate the capability of the method of Figure 1, the thermal resistances of two different thick silicon substrates (520μm thick and 370μm thick) are measured and the results are shown in Table 2. The thermal resistance of 520μm thick silicon is measured to be 0.036 ± 0.003 C cm^2/W and it is in reasonable agreement with the value of 0.035 C cm^2/W which is calculated by using the thermal conductivity of silicon (148W/mK). The thermal resistance of 370μm thick silicon is measured to be 0.026 ± 0.003 C cm^2/W and it is also in reasonable agreement with the value of 0.025 C cm^2/W which is calculated by using the thermal conductivity.

$$\text{Thermal resistance of a sample (C cm}^2\text{/W)} = \frac{T_1 - T_2\ (\text{C})}{\text{Input Power (W)}} - (2 \times \text{Contact resistance} + \text{Thermal resistance of copper blocks})$$

Equation 1. Derivation of the thermal resistance of a sample.

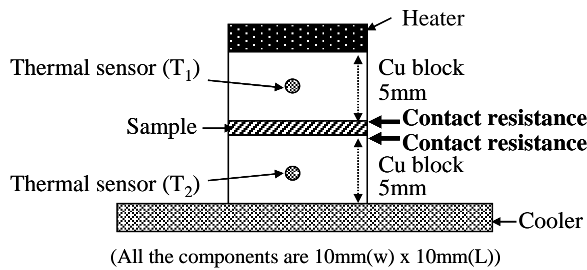


Figure 1. Steady-state thermal resistance measurement method (cross sectional view)

Contact material	Contact resistance (C cm^2/W)	Fluctuation (C cm^2/W)
Grease	0.19	± 0.01
Liquid metal	**0.008**	**± 0.001**

TABLE 1: Comparison of the contact resistance and its resultant fluctuation between grease and liquid metal, which are measured by the method of Figure 1.

Μεασυρεδ ρεσυλτ οφ 520μm thick

	Thermal resistance (C cm^2/W)
Calculated (from thermal conductivity)	0.035
Measured	**0.033 - 0.039**

Measured result of 370μm thick silicon

	Thermal resistance (C cm^2/W)
Calculated (from thermal conductivity)	0.025
Measured	**0.023 - 0.029**

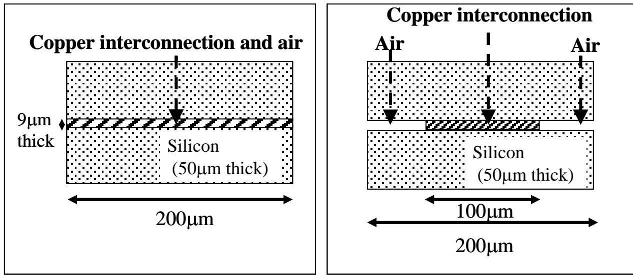
TABLE 2: Thermal resistance measurement results of 520μm thick silicon and 370μm thick silicon substrates.

3. Results and discussions

3.1. Thermal resistance of a silicon substrate

The thermal resistance of a silicon substrate is simulated (finite-element-method) employing a μm-scale model which is depicted in the right of Figure 2A. The comparison of a horizontally continuum model and a μm-scale model (this study) is shown in Figure 2A. The simulated temperature distributions of both models, when horizontally uniform heat is applied at the bottom, are also shown in Figure 2B. In both models, a 9μm thick copper interconnection is sandwiched by two 50μm thick silicon substrates. The diameter of a copper interconnection is 100μm and the interconnection pitch is 200μm. In a horizontally continuum model, the thermal conductivity is assumed to be horizontally uniform and the unified thermal conductivity of a copper interconnection and air is assigned to the whole of a copper interconnection and air. The unified thermal conductivity is derived based on the area ratio between a copper interconnection and air. In this case, the area ratio of a copper interconnection is 0.196 and that of air is 0.804. The thermal conductivity of copper (398W/mK) is multiplied by 0.196 and that of air (0.026W/mK) is multiplied by 0.804. The addition of these two equals the unified thermal conductivity. Compared with a horizontally continuum model, in a μm-scale model, the thermal conductivity of copper is assigned to a μm-scale copper interconnection and that of air is assigned to air. As depicted in Figure 2B, the temperature distribution of a μm-scale model is simulated to be larger than that of a horizontally continuum model (which means that the thermal resistance by a μm-scale model is simulated to be larger than that by a horizontally continuum model), because in a μm-scale model, thermal conduction path is limited by a μm-scale interconnection and the area of the effective thermal conduction path is smaller. Simulated thermal resistances by both models are described in Table 3. The simulated thermal resistance by a horizontally continuum model is 0.008 C cm^2/W and that by a μm-scale model is 0.019 C cm^2/W, which indicates that the later is more than twice of the former. The difference between the former and the later is dependent on the interconnection diameter, pitch and silicon thickness. A μm-scale model requires more time to be built than a horizontally continuum model, however a μm-scale model gives more precise results because it reproduces an actual structure, therefore a μm-scale model is employed in this study. When multi-chips are stacked, an error of the thermal resistance of a single-stacked-chip adds up to a large amount, and it results in misunderstanding the thermal resistance of 3D chip stack. In this regard, a μm-scale model is essential.

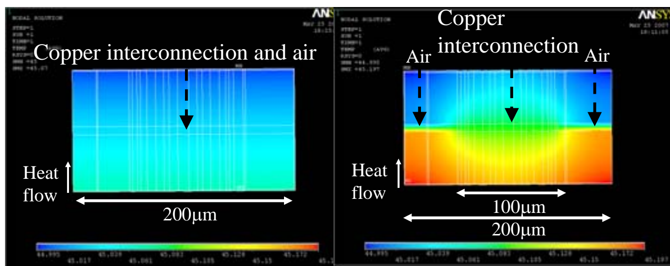
By using a μm-scale model (in the right of Figure 2A), the thermal resistance of a silicon substrate with various interconnection diameters and pitches is simulated and the dependence of the simulated thermal resistance on the interconnection diameter and pitch is clarified. It is obtained by changing the interconnection diameter and pitch in a μm-scale model. The thermal resistance of a silicon substrate is derived by the temperature difference between a copper interconnection and the top or the bottom of the model (the top of the upper silicon substrate or the bottom of the lower silicon substrate.) The dependence of the simulated thermal resistance of a silicon substrate on the interconnection diameter and pitch is described in Table 4. As can be seen in Table 4, even though the area ratio of an interconnection (i.e. the interconnection area divided by the total chip area) is the same in 20μm pitch, 50μm pitch, 100μm pitch and 200μm pitch, as the pitch increases, the simulated thermal resistance of a silicon substrate gradually increases. When the



Horizontally continuum model

µm-scale model (This study)

Figure 2A. A horizontally continuum model and a µm-scale model. (cross sectional view) (The thermal conductivity of silicon is set to be 148W/mK, that of copper is 398W/mK, and that of air is 0.026W/mK, respectively.)



Horizontally continuum model

µm-scale model (This study)

Figure 2B. Simulated temperature distribution of a horizontally continuum model and that of a µm-scale model. (cross sectional view)

	Simulated thermal resistance (C cm²/W)
Horizontally continuum model	**0.008**
µm-scale model	**0.019**

TABLE 3: Simulated thermal resistance of a horizontally continuum model and that of a µm-scale model

interconnection diameter is the same (200µm pitch and 500µm pitch), as the pitch increases, the simulated thermal resistance of a silicon substrate significantly increases.

Further, the thermal resistance of a silicon substrate which is sandwiched by two interconnections is simulated by using a µm-scale model which is described in Figure 3. In this case, the thermal resistance of a silicon substrate is derived by the temperature difference between two copper interconnections which sandwich a silicon substrate. The dependence of the simulated thermal resistance on the interconnection diameter and pitch is depicted in Table 5.

Interconnection pitch (interconnection dia.)	20µm (10µm dia.)	50µm (25µm dia.)	100µm (50µm dia.)	200µm (100µm dia.)	500µm (100µm dia.)
Simulated thermal resistance of a silicon substrate (C cm²/W)	0.004	0.005	0.005	0.007	0.040

TABLE 4: Dependence of the simulated thermal resistance of a silicon substrate on the interconnection diameter and pitch (the thickness of a silicon substrate is 50µm and the thermal resistance of a bare 50µm thick silicon substrate is calculated to be 0.0034 C cm²/W.)

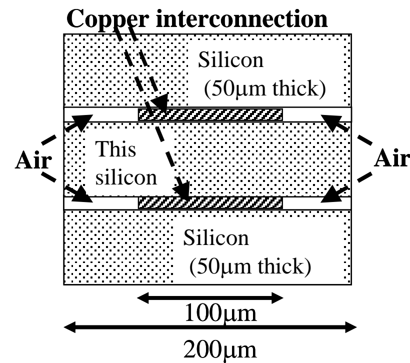


Figure 3. µm-scale model of a silicon substrate which is sandwiched by two interconnections. (cross sectional view)

Interconnection pitch (interconnection dia.)	20µm (10µm dia.)	50µm (25µm dia.)	100µm (50µm dia.)	200µm (100µm dia.)	500µm (100µm dia.)
Simulated thermal resistance of a silicon substrate (C cm²/W)	0.004	0.006	0.007	0.012	0.061

TABLE 5: Dependence of the simulated thermal resistance of a silicon substrate on the interconnection diameter and pitch, when it is sandwiched by two interconnections.

The simulated thermal resistances of a silicon substrate when it is sandwiched by two interconnections (Table 5.) are larger than those when one interconnection is sandwiched by two silicon substrates (Table 4.). This is because in the former (Table 5.), the area of the effective thermal conduction path is more restricted as a result of two interconnections. The difference between the former and the later becomes significant at large interconnection pitches. Silicon thickness

(in this case, 50μm thick) influences the dependence of this difference on the interconnection diameter and pitch.

3.2. Thermal resistance of an interconnection

The thermal resistance of an interconnection can be experimentally obtained by using the thermal resistance measurement method which is shown in Figure 1. and the simulated thermal resistance of a silicon substrate which is described in 3.1. Sakuma et al.[15] investigated copper-tin (CuSn) interconnections, including the mechanical strength and electrical resistance. The interconnection sample in this study is composed of 200μm pitch (100μm diameter) and 9μm thick copper-tin (CuSn) interconnections[15] sandwiched by two 730μm thick silicon substrates. Firstly, the thermal resistance of the interconnection sample is measured by the thermal resistance measurement method of Figure 1. Secondly, twice of the simulated thermal resistance of a 730μm thick silicon substrate with a 200μm pitch (100μm diameter) interconnection, is subtracted from the measured thermal resistance. The thermal resistance of an interconnection can be obtained in this manner and it contains the contact thermal resistance between a silicon substrate and an interconnection, the contact thermal resistance between metals of an interconnection. Off-course it also includes the thermal resistances of intermetallic compounds, such as Cu_3Sn. The thermal resistance of a 200μm pitch (100μm diameter), 9μm thick CuSn interconnection is obtained to be 0.078 ± 0.018 C cm^2/W.

From the thermal conductivity of copper (398W/mK), that of tin (67W/mK), and that of nickel (91W/mK), the thermal resistance of a 200μm pitch, 9μm thick CuSn interconnection can simply be calculated to be 0.003 C cm^2/W. In this calculation, intermetallic compounds and very thin gold (Au) and titianium (Ti) are not considered. The difference between the experimental thermal resistance of 0.078 ± 0.018 C cm^2/W and the calculated thermal resistance of 0.003 C cm^2/W arises from the contact thermal resistance between a silicon substrate and an interconnection, the contact thermal resistance between metals of an interconnection, and the thermal resistance of intermetallic compounds, such as Cu_3Sn. The thermal conductivities of intermetallic compounds were studied by Terada et al.[16] and Jacobsson et al.[17], and it was demonstrated that the thermal conductivities of intermetallic compounds are dependent whether they are in the ordered crystal phase or they are in the disordered phase and those in the ordered crystal phase are higher than those in the disordered phase. In this study, the phases of intermetallic compounds in a copper-tin interconnection are not yet analyzed and it is a subject for the future study. With regard to the contact thermal resistance, Bai et al.[18] measured the contact thermal resistance between eutectic lead-tin solder and copper by laser-flash method to be from 0.011 to 0.033 C cm^2/W. Although the metal compositions in Bai's study are not the same as those in this study, it is pointed out that the difference between the experimental thermal resistance of 0.078 ± 0.018 C cm^2/W and the calculated thermal resistance of 0.003 C cm^2/W in this study is consistent with the measured contact thermal resistance by Bai et al. (from 0.011 to 0.033 C cm^2/W).

When the thermal resistance of a 200μm pitch , 9μm thick CuSn interconnection obtained in this study (0.078 ± 0.018 C cm^2/W) is transformed into thermal conductivity, it is 6.24 ± 1.41 W/mK. Shimada et al.[19] measured the thermal resistances of B^2it (Buried Bump Interconnection technology) samples and then derived the thermal conductivity of an interconnection of B^2it samples. An interconnection of B^2it was made of silver paste and its thermal conductivity was derived to be 10 W/mK. Experimental thermal conductivity of a 200μm pitch, 9μm thick CuSn interconnection in this study (6.24 ± 1.41 W/mK) exceeds that of grease (at best around 5 W/mk), but does not reach that of an interconnection of B^2it samples (10 W/mK). It is expected that a 200μm pitch, 9μm thick CuSn interconnection in this study work for thermal conduction more than grease does, but not as much as an interconnection of B^2it samples does.

3.3. Thermal resistance of back-end-of-the-line (BEOL)

The thermal conductivities of interlayer dielectrics (ILDs) of BEOL are supposed to be low and the thermal resistance of BEOL is considered to influence the thermal characteristics of 3D chip stack. Therefore, the thermal conductivities of ILDs of BEOL with the 45nm technology node are experimentally measured by the methods described in reference 20[20] and the measured effective thermal conductivities are described in Table 6. A simple simulation model of the 45nm BEOL[21,22] is built as shown in Figure 4. and by assigning the experimental thermal conductivities to ILDs of the model, the thermal resistance of the 45nm BEOL is simulated to be 0.003 C cm^2/W. This thermal resistance is derived by the difference between the average temperature at the top surface and that at the bottom surface. In this model, only vertical lines are considered (horizontal lines are not considered) and the contact thermal resistance of copper lines between levels is not considered. Furthermore, the thermal resistance of BEOL is supposed to be affected with though-silicon-via(TSV), which is not included in this study. It should be noted that the thermal resistance of BEOL, considering horizontal lines and the contact thermal resistance of copper lines and including the effect of though-silicon-via(TSV), needs to be further evaluated in the future study.

Level	Keff(W/mK)
L1	1.07
L2	0.65
L3	0.64
L4	0.499
L5	0.503
L6	0.590
L7	0.565
L8	0.751

Table 6 : Measured effective thermal conductivities of ILDs of the 45nm BEOL ('L' means level and 'Keff' means effective thermal conductivity,

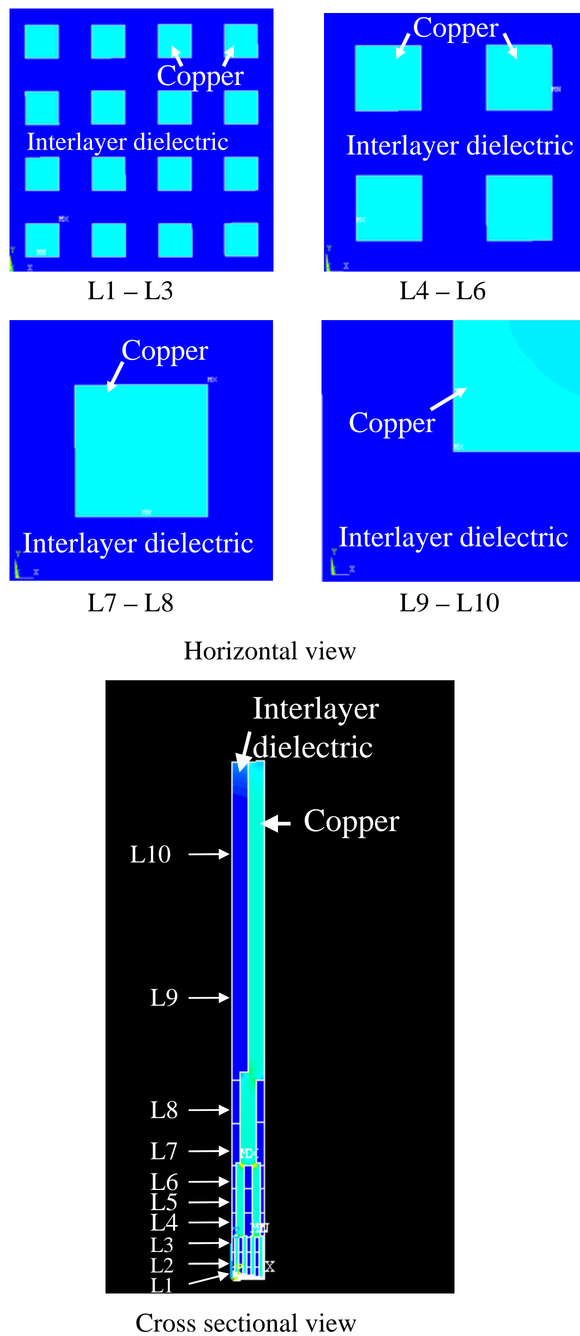


Figure 4. 45nm BEOL model ('L' means level.)

3.4. Total thermal resistance of 3D chip

Based on the thermal resistances of the silicon substrate (3.1), the interconnection (3.2) and the BEOL (3.3), the total thermal resistance of 3D chip stack whose configuration is shown in Figure 5 is estimated. The 3D chip stack in Figure 5 is composed of four silicon substrates, three BEOLs and three interconnections. The total thermal resistance of 3D chip stack is derived by multiplying the thermal resistance of the silicon substrate (Table 4.) by 2, multiplying the thermal resistance of the silicon substrate when it is sandwiched by two interconnections (Table 5.) by 2, multiplying the thermal resistance of the interconnection by 3, multiplying the thermal resistance of the 45nm BEOL by 3, and adding these. The thermal resistance of an interconnection in the interconnection pitches of 20μm, 50μm and 100μm, is assumed to be equal to the thermal resistance of a 200μm pitch, 9μm thick CuSn interconnection obtained in the section of 3.2, because the area ratio of an interconnection (i.e. the interconnection area divided by the total chip area) in these interconnection pitches is the same. (When the interconnection pitch is 20μm, 50μm and 100μm, the interconnection diameter is set to be 10μm, 25μm and 50μm, respectively.) This assumption is based on that the thermal resistance of an interconnection consists of the thermal resistance of metals (including intermetallic compounds), the contact thermal resistance between a silicon substrate and an interconnection, and the contact thermal resistance between metals of an interconnection. When the interconnection pitch is 500μm, the interconnection diameter is set to be 100μm. The thermal resistance of a 500μm pitch interconnection is calculated by considering the area ratio of an interconnection as follows. When the interconnection pitch is 200μm (the interconnection diameter is 100μm) and 500μm (the interconnection diameter is 100μm), the area ratio of an interconnection is 0.196 and 00314, respectively. The thermal resistance of a 200μm pitch (100μm diameter) interconnection obtained in this study (0.078 ± 0.018 C cm^2/W) is multiplied by 0.196/0.0314 and it is calculated to be 0.487 ± 0.112 C cm^2/W. The thermal resistance of a 500μm pitch (100μm diameter) interconnection is derived in this manner. With regard to the thermal resistance of the BEOL, the thermal resistance increase because of limited thermal conduction path by a μm-scale interconnection is not considered.

The dependence of the estimated total thermal resistance on the interconnection pitch and diameter is depicted in Figure 6A. For example, when the interconnection pitch is 20μm, the total thermal resistance is estimated to be 0.259 ± 0.054 C cm^2/W. When the interconnection pitch is below 200μm, the estimated total thermal resistance is almost invariable around 0.025 - 0.029 C cm^2/W. It is because the thermal resistance of interconnections is dominant and the thermal resistances of silicon substrates and BEOLs are smaller than that of interconnections. As the interconnection pitch increases to 500μm (the interconnection diameter is 100μm), the estimated total thermal resistance significantly increases because of mainly the thermal resistance of interconnections (1.461 ± 0.336 C cm^2/W) and also that of silicon substrates (0.202 C cm^2/W). These results indicate that the thermal resistance of interconnections is the bottleneck in the total thermal resistance of 3D chip stack.

Colgan et al. in IBM T. J. Watson Research Center developed silicon micro channel cooler[23] and it is adopted here as an example of cooling methods. They estimated the thermal resistances of TIM (indium), base and fins of silicon microchannel cooler to be totally 0.08 C cm^2/W. Combined with it, the thermal resistance from the bottom of 3D chip stack to the water inlet, when the interconnection pitch is 20μm as an example, is calculated to be 0.339 ± 0.054 C

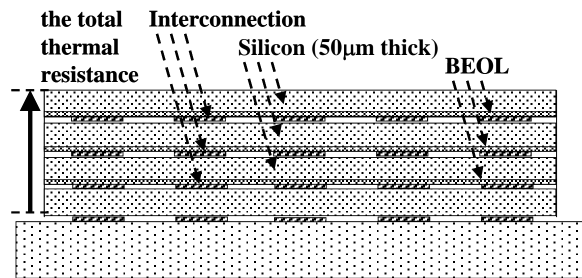


Figure 5. Present 3D chip stack model (The total thermal resistance is composed the thermal resistances of four silicon substrates, those of three BEOLs and those of three interconnections.)

cm^2/W (0.08 + (0.259 ± 0.054) C cm^2/W). They also assumed the maximum junction temperature and the water inlet temperature to be 85 and 22 C, respectively, and then ΔT (the maximum junction temperature minus the water inlet temperature) is 63 C. When ΔT is 63 C, maximum allowable power density (maximum allowable power density here corresponds to just the maximum allowable horizontally-uniform power density at the bottom of 3D chip stack.) is calculated to be 191 ± 30 W/cm^2 (63 C / (0.339 ± 0.054) C cm^2/W). The dependence of the estimated maximum allowable power density on the interconnection pitch and diameter is depicted in Figure 6B. As similar to Figure 6A, when the interconnection pitch is below 200μm, the estimated maximum allowable power density is close within 150 - 220 W/cm^2. When the interconnection pitch increases to 500μm (the interconnection diameter is 100μm), the maximum allowable power density significantly decreases to 37.3 ± 7.2 W/cm^2.

According to International Technology Roadmap for Semiconductors (ITRS) 2006, cost-performance MPU maximum power density at 2018 (10 years from now) is predicted to be 108 W/cm^2. When 3D chip stack contains additional heat-dissipating chips other than a MPU, the total power density of 3D chip stack increases. The above maximum allowable power density at the bottom of 3D chip stack of 191 ± 30 W/cm^2 is an example when the configuration of Figure 5, the interconnection pitch of 20μm, and silicon microchannel cooler is assumed, however it clearly demonstrates that the optimizations of the configuration of 3D chip stack, the interconnection pitch, a cooling method, are important to suffice the thermal requirement of 3D chip stack. Figure 6B implies that when 3D chip stack contains more than two MPUs (the power density is more than 216 W/cm^2) in the configuration of Figure 5, even when silicon microchannel cooler is adopted and the interconnection pitch is 20μm, additional cooling solutions are considered to be necessary in order to meet the thermal requirement.

It should be noted that hot-spots are not considered in this study. For investigating the thermal management of 3D chip stack in the future study, hot-spots locations should be taken into consideration to evaluate the effect of cooling solutions.

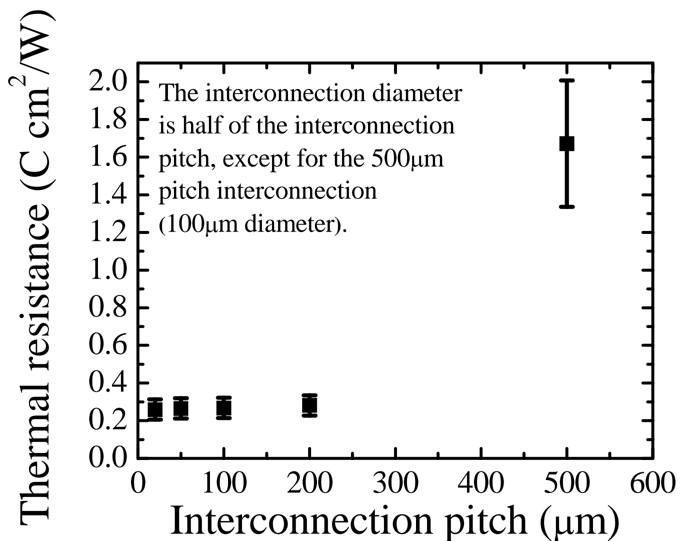


Figure 6A. Dependence of the estimated total thermal resistance of 3D chip stack (Figure 5) on the interconnection pitch and diameter.

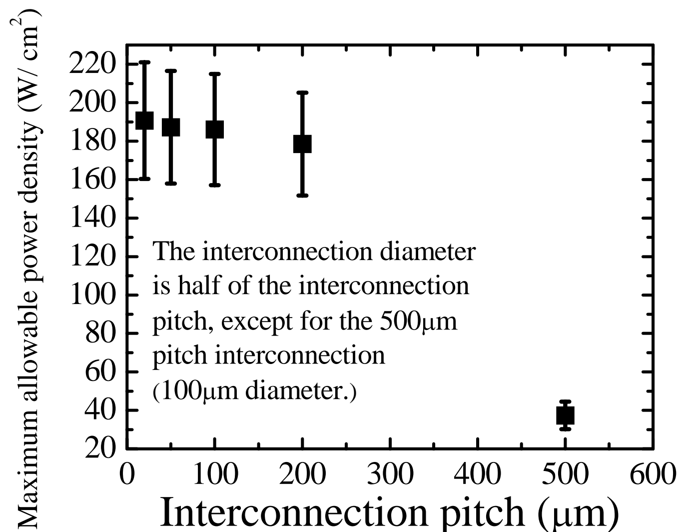


Figure 6B. Dependence of the estimated maximum allowable power density at the bottom of the 3D chip stack (Figure 5) on the interconnection pitch and diameter.

4. Concusions

The thermal resistance of 3D chip stack is investigated from the thermal resistance of each component of a single-stacked-chip. Firstly, the thermal resistance of a silicon substrate with various interconnection pitches is simulated by employing a μm-scale model, and the dependence of the simulated thermal resistance on the interconnection pitch and diameter is presented. Secondly, steady-state thermal resistance measurement method using liquid metal as contact material is achieved, and the thermal resistance of a copper-tin interconnection is obtained by the combination of measurement and simulation. Thirdly, the thermal resistance

of BEOL of the 45nm technology node is simulated by building an actual structural model and by assigning the experimental thermal conductivities to interlayer dielectrics of the model. The total thermal resistance of 3D chip stack with the configuration of four-stacked chips is estimated based on the thermal resistances of the silicon substrate, the interconnection and the 45nm BEOL. The dependence of the estimated total thermal resistance of the 3D chip stack on the interconnection pitch and diameter is clarified. Silicon microchannel cooler is adopted as an example of cooling methods and it is estimated how much power density of the 3D chip stack is allowed as a function of the interconnection pitch and diameter. By comparing with the ITRS prediction of cost-performance MPU maximum power density at 2018, it is demonstrated that the optimizations of the configuration, the interconnection pitch, cooling methods and solutions are critical to satisfy the thermal requirement of 3D chip stack.

Acknowledgments

I would like to acknowledge Dr. E. G. Colgan, Dr. P. S. Andry and Dr. J. H. Magerlein in IBM T. J. Watson Research Center, Dr. R. J. Bezama in IBM Systems & Technology Group, for the very informative discussions. I would also like to acknowledge Dr. Y. Takahashi in IBM Integrated Supply Chain for the support on simulation, and Dr. S. Kohara in IBM Tokyo Research Laboratory for the support on measurement.

References

[1] Chiang, T.-Y.; Souri, S. J.; Chui, C. O.; Saraswat, K. C. “Thermal analysis of heterogeneous 3-D ICs with various integration scenarios ”, *International Electron Devices Meeting (IEDM),* **2001**, p.681.

[2] Im, M.; Banerjee, K. “Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs”, *International Electron Devices Meeting (IEDM),* **2000,** p.727.

[3] Puttaswamy, K.; Loh, G. H. “Thermal analysis of a 3D die-stacked high-performance icroprocessors”, *Great Lakes Symposium on VLSI,* **2006**, p.19.

[4] Puttaswamy, K.; Loh, G. H. “Thermal herding: Microarchitecture techniques for controlling hotspots in high-performance 3D-integrated processors”, 13th IEEE International Symposium on High Performance Computer Architecture *(HPCA)*, **2007,** p.193.

[5] Das, S.; Chandrakasan A.; Reif, R. “Timing, energy and thermal performance of three-dimensional integrated circuits”, *Great Lakes Symposium on VLSI,* **2004,** p.338.

[6] Hua, H.; Mineo, C.; Schoenfliess, K.; Sule, A.; Melamed, A.; Jenkai R.; Davis, W. R. “Exploring compromises among timig, power and temperature in three-dimensional integrated circuits ”, *DAC,* **2006,** p.997.

[7] Chiang, T.-Y.; Banerjee, K.; Saraswat, K. C. "Effect of via separation and low-k dielectric materials on the thermal characteristics of Cu Interconnects", *International Electron Devices Meeting (IEDM),* **2000**, p.261.

[8] Chiang, T.-Y.; Banerjee, K.; Saraswat, K. C. "Analytical thermal model for multilevel VLSI interconnects incorporating via effect", *IEEE ELECTRON DEVICE LETTERS,* **2002,** vol.23, p.31.

[9] Chiang, T.-Y.; Saraswat, K. C."Impact of vias on the thermal effect of deep sub-micron Cu/low-k interconnects", *Symposium on VLSl Technology Digest of Technical Papers,* **2001,** p.141.

[10] Goplen, B.; Sapatnekar, S. ”Thermal via placement in 3D ICs”, *ISPD,* **2005,** p.167.

[11] Yu, H.; Shi, Y.; He, L.; Karnik, T. “Thermal via allocation for 3D ICs considering temporally and spatially variant thermal power”, *ISLPED,* **2006,** p.156.

[12] Yamaji, Y.; Ando, T.; Morifuji, T.; Tomisaka, M.; Sunohara, M.; Sato, T.; Takahahi. K. ”Thermal characterization of bare-die stacked module with Cu thorough-vias”, *Electronic Components and Technology Conferences*, **2001**, p.730.

[13] Sweet, J. N.; Peterson, D. W.; Chu, D.; Bainbridge, B. L.; Gassman, R. A.; Reber, C. A. ”Analysis and measurement of thermal resistance in a 3-dimensional silicon multichip module populated with assembly test chips”, *9th Annual IEEE Semiconductor Thermal Measurement and Management Symposium,* **1993,** p.1.

[14] Gurrum, S.; Suman, S.; Joshi, Y.; Fedorov, A.; “Thermal issued in next-generation integrated circuits,” *IEEE Transactions on Device and Materials Reliability,* **2004,** vol. 4, p.709.

[15] Sakuma, K.; Andry, P. S.; Dang, B.; Maria, J.; Tsang, C. K.; Patel, C.; Wright, S. L.; Webb, B.; Sprogis, E.; Kang, S. K.; Polastre, R.; Horton, R.; Knickerbocker, J. U. “3D chip stacking technology with low-volume lead-free interconnections”, *Electronic Components and Technology Conferences*, **2007**, p.627.

[16] Terada Y.; Ohkubo, K.; Mohri, T.; Suzuki, T. “Thermal conductivity of intermetallic compounds with metallic bonds,” *Materials Transactions*, **2002,** vol.43, p.3167.

[17] Jacobsson, P.; Sundqvist, B. “Pressure dependence of the thermal and electrical conductivities of the intermetallic compounds AnCu ans AuCu3,” *J. Phys. Chem. Solids,* **1998,** vol.49, p.441.

[18] Bai, J.G.; Zhang, Z. Z.; Lu, G.-Q.; Hasselman, D. P. H. “Measurement of solder/copper interfacial thermal resistance by the flash technique,” *Internatinal Journal of Thermophysics*, **2005,** vol.26, p.1607.

[19] Shimada, O.; Hisano, K.; Iwasaki, H.; Ishizuka, M.; Fukuoka, Y. “Thermal management estimations for Buried Bump Interconnection Technology printed wiring boards with bump (filled via) interconnection”, *Inter Society Conference on Thermal Phenomena,* **1998,** p.468.

[20] F. Chen, J. Gill, D. Harmon, T. Sullivan, B. Li, A. Strong, H. Rathore, D. Edelstein, C-C, Yang, A. Cowley, and L. Clevenger, **“**Measurements of effective thermal conductivity for advanced interconnect structures with various composite low-k dielectrics”, *International Reliability Physics Symposium (IRPS),* **2004,** p.68.

[21] Narasimha, S.; Onishi, K.; Nayfeh, H. M.; Waite, A.; Weybright, M.; Johnson, J; Fonseca, C.; Corliss, D.; Robinson, C.; Crouse, M.; Yang, D.; Wu, C-H. J.; Gabor, A.; Adam, T.; Ahsan, I. et al. "High-performance 45nm SOI technology with enhanced strain, porous low-k BEOL, and immersion lithography", *International Electron Devices Meeting (IEDM)*, **2006**.

[22] Sankaran, S.; Arai, S.; Augur, R.; Beck, M.; Biery, G.; Bolom, T.; Bonilla, G.; Bravo, O.; Chanda, K.; Chae, M.; Chen, F.; Clevenger, L.; Cohen, S.; Cowley, A.; Davis, P. et al. "A 45 nm CMOS node Cu/Low-k/ Ultra Low-k PECVD SiCOH (k=2.4) BEOL Technology", *International Electron Devices Meeting (IEDM)*, **2006**.

[23] Colgan, E. G.; Furman, B.; Gaynes, M.; Graham, W.; LaBianca, N.; Magerlein, J. H.; Polastre, R. J.; Rothwell, M. B.; Bezama, R. J.; Choudhary, R.; Marson, K.; Toy, H.; Wakil, J.; Zitz, J; Schmidt, R. "Practical implementation of silicon microchannel coolers for high power chips", *21st Annual IEEE Semiconductor Thermal Measurement and Management Symposium,* **2005***,* p.1.