Microelectronics Reliability xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Microelectronics Reliability

journal homepage: www.elsevier.com/locate/microrel

Thermal impact of extreme die thinning in bump-bonded three-dimensional integrated circuits

Samson Melamed *, Naoya Watanabe, Shunsuke Nemoto, Haruo Shimamoto, Katsuya Kikuchi, Masahiro Aoyagi

National Institute of Advanced Industrial Science and Technology (AIST), Central 1, 1-1-1 Umezono, Tsukuba 305-8560, Japan

ARTICLE INFO

Article history:
Received 6 March 2017
Received in revised form 23 May 2017
Accepted 23 May 2017
Available online xxxx

Keywords:
3DIC
Die thinning
Thermal resistance
Integrated circuit

ABSTRACT

In three-dimensional integrated circuits (3DICs) aggressive wafer-thinning can lead to large thermal gradients, including spikes in individual device temperatures. In a non-thinned circuit, the large bulk silicon wafer on which devices are built works as a very good thermal conductor, enabling heat to diffuse laterally. In this paper we use measurement and simulation results to examine the normalized temperature rise of an on-chip heater in a two-tier bump-bonded 3D stacked system. We begin by experimentally validating our simulation model and then use it to formulate best and worst case scenarios for the temperature rise in such a system. Die thinning is seen to have a pronounced effect on the thermal response, which can adversely affect system reliability. Thinning the top tier from 725 μm to 10 μm resulted in approximately an 8 times increase in the normalized temperature rise of the heater in our test chip for the worst case scenario and just over a 6 times increase for the best case scenario.

1. Introduction

In three-dimensional integrated circuits (3DICs) aggressive wafer-thinning can lead to large thermal gradients [1], including spikes in individual device temperatures [2]. In a non-thinned circuit, the large bulk silicon wafer on which devices are built works as a very good thermal conductor, enabling heat to diffuse laterally. When the wafer is thinned, the heat generated by individual devices can no longer spread out as easily, with the result being "hotspots" in the temperature profile. In [2] it was found that single clock buffers could create hotspots when they were located in the upper tiers of a fully-depleted silicon-on-insulator (FDSOI) 3DIC. Torregiani et al. have done a simulation-based study on the impact of various parameters on the size and temperature of dies as thin as 20 μm [3]. Santos et al. simulated the effects of substrate thickness and die bonding options on a memory-on-logic 3D circuit and found increased temperatures at decreased thicknesses [4]. Kato et al. investigated how thinning the die results in increased hotspot temperature [5]. The effects of high temperatures on semiconductor devices include reduced carrier mobility, increased subthreshold leakage [6], and decreased mean time to failure due to electromigration [7]. Increased temperatures not only directly affect system reliability, but also makes it increasingly challenging to design working circuits: changes in device performance with temperature necessitates the consideration of on-chip temperature variations when attempting to obtain timing closure in digital circuits [8,9].

In a bump-bonded system, two or more wafers or dies are vertically stacked. Through silicon vias (TSVs) are placed into thinned upper tiers to allow for vertical electrical connections. Bump bonds are then used to connect between a lower tier's electrical wiring and an upper tier's TSV. The gap caused by these bump bonds must be filled with an underfill material, which typically has a poor thermal conductivity around two orders of magnitude lower than that of silicon.

TSVs are typically limited to an aspect ratio (depth to diameter) of approximately 10 or less in high volume production. For instance, at a wafer thickness of 20 μm, the TSVs will need to have a diameter of approximately 2 μm or more. This is significantly larger than a traditional back end of line (BEOL) via, meaning that routing vertically between tiers comes with a large overhead in area. There is, therefore, a significant interest in thinning the tiers that contain TSVs so that the TSVs can be sized more similarly to traditional vias. The downside to this, however, is that thinning the tiers increases the lateral thermal resistance of the individual tiers in the system by removing much of the bulk silicon.

The handling of thinned wafers is significantly challenging. Die thinning can be accomplished on fully fabricated wafers by first mounting them on a support glass substrate. This allows for the sample to stay rigid during thinning, and allows for additional steps, such as microbumping and dicing to be performed.

Oprins et al. have previously conducted measurements to investigate temperature changes in hotspots caused by varying densities of TSVs in a system with thinned wafers. The wafers were hybrid-bonded with Cu—Cu bonds and a thin dielectric layer [10]. In a prior work we experimentally confirmed the normalized temperature increase

* Corresponding author.
E-mail address: samson.melamed@aist.go.jp (S. Melamed).

http://dx.doi.org/10.1016/j.microrel.2017.05.030
0026-2714/

(temperature increase divided by power) for a heater in a simplified stackup that was designed to replicate the thermal environment of a bump-bonded system [11]. This simplified stackup was formulated as a general worst-case scenario. In this work we begin by experimentally validating computer-based simulation results by comparing them to measurements. Next, we use our validated thermal simulation model to study additional cases, including more aggressively thinned circuits with wafer thicknesses as low as 10 μm. We further add simulations where the bottom low-conductivity substrate is replaced with high-conductivity material, enabling a comparison of both a general worst-case scenario as well as a best-case scenario for the structure. In all cases we did not include individual TSVs and bumps, as their location is design-dependent. By not including these layout-dependent objects, we are able to provide broad insight on these structures without being tied to a specific process or design. In this paper we present results for the normalized temperature increase for the heater with upper die thicknesses from 725 μm (unthinned) to 10 μm and a comparison of the best and worst case scenarios for temperature rise.

The remainder of this paper is organized as follows. In Section 2 we discuss bump-bonded systems and our simplified structure for recreating their thermal environment. In Section 3 we describe the heater chip we designed along with our simulation model. In Section 4 we present the simulation results, compare them to measurements, and describe the thermal effects of die thinning.

2. Background

A sample bump-bonded stack is shown in Fig. 1(a), with a heater chip embedded into the top silicon chip. The system is composed of two tiers. The top tier will need to be thinned in order to insert through silicon vias (TSVs) to provide for electrical connections between the top and bottom tiers. When the top silicon chip is thick, heat from the heater can easily spread out laterally, and then pass through the underfill/bump joint layer in order to reach the heatsink. The underfill/bump joint layer typically has a thermal conductivity approximately two orders of magnitude less than that of silicon. When the heat can spread out laterally first, the heat can then pass through the underfill/bump region in parallel over a wide area, which presents a relatively low thermal resistance. When the upper silicon wafer is thinned, however, there is no longer a low resistivity path for the heat to initially spread laterally: the thinned wafer has a much higher lateral resistance than that of the unthinned wafer. This is most easily understood by looking at the equation for the thermal resistance of a simple block of material:

$$R_\theta = \frac{l}{A \times k} \tag{1}$$

where l is the length of the block on the path parallel to the heat flow (m), A is the cross-sectional area of the block perpendicular to the path of the heat flow (m^2) and k is the thermal conductivity of the block (W/m·K). For a thin slice of the system of Fig. 1(a) we find that as the upper die is thinned the cross-sectional area from heater laterally along the upper Si chip decreases, indicating an increased resistance to lateral spreading. Should the underfill/bump joint layer have a similar thermal conductivity to silicon, this would not be an issue, as the heat could immediately move through the underfill/bump joint layer and continue spreading in the bottom chip. Since the underfill/bump joint layer has a high resistance, we expect to see that thinning the top tier will significantly increase the thermal resistance from the heater to the heatsink.

3. Methodology

In this section we first describe the simplified stackup that is used to replicate the thermal environment of the system in Fig. 1(a). We then describe the test chip that we have fabricated to experimentally determine the thermal resistance from the top heater to the heatsink for various thicknesses of the upper Si chip. We finally describe the simulation model used to perform an accurate simulation in Keysight HeatWave.

3.1. Simplified model

In this section we will begin with a brief review of our approach to create a simplified model of the system shown in Fig. 1(a).

Our simplified model is created by removing TSVs, and replacing the underfill and bump bonds with a die attach film, which serves a similar role by provided a low-conductivity barrier. Removing the TSVs and bumps creates a worst-case thermal environment to aid in investigating the effects of thinning, while still staying true to the thermal properties of the non-simplified system.

Our simplified system is shown in Fig. 1(b). This simplified model provides a suitable structure to enable the creation and measurement of a hotspot for the purpose of comparing measurement results to a computer-based simulation model, which is the first target of this paper. Additional details on our modeling approach can be found in [11].

3.2. Test chip

To experimentally test the effects of die thinning in a bump-bonded 3D stacked system, we fabricated heater chips with diodes to measure temperatures at multiple locations (see Fig. 2). A simplified version of the chip layout is shown in Fig. 3(a). A central 100 μm × 100 μm resistive heater area is located in the middle of the 2 mm × 2 mm chip to generate a hotspot. The heater itself is formed from five 4 μm × 100 μm n + diffusion resistors connected in parallel, and can also be used to measure temperature. Four diodes are then located at intervals of 25 μm from the bottom side of the heater to allow for electrical measurement of the chip temperature at various points away from the central heater. The wafers containing the test chip were diced to 10 mm × 11 mm as shown in Fig. 3(b) to allow for improved handling during backgrinding to thin the silicon substrate. The chips had a silicon substrate thickness of h = 725 μm. These chips were then thinned to a variety of silicon substrate thicknesses between h = 725 μm and h = 20 μm.

Quartz wafers of thickness 400 μm were separately prepared. The wafers were laminated onto 20 μm thick (before oven cure) version of Hitachi Chemical's HR-5104 die attach film, which comes as a sheet

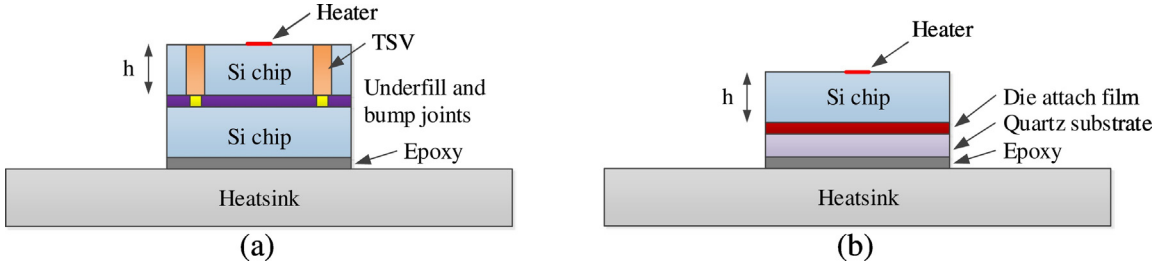

Fig. 1. (a) Two-tier stacked system with underfill and bump joints. (b) Simplified structure for thermal measurements including a layer with poor thermal conductivity (die attach film) under the top chip to mimic the thermal properties of underfill and bump joints as well as a quartz substrate to increase temperatures to a range that is more easily measurable.

Please cite this article as: S. Melamed, et al., Thermal impact of extreme die thinning in bump-bonded three-dimensional integrated circuits, Microelectronics Reliability (2017), http://dx.doi.org/10.1016/j.microrel.2017.05.030