Motivations for three-dimensional (3D) integration include reduction in system size, interconnect delay, power dissipation and enabling hyper-integration of chips fabricated using disparate process technologies. Although various low-power commercial products exploit the advantages of improved performance and increased device packing density realized by 3D stacking of chips (using wirebonds), such technologies are not suitable for high-performance chips due to ineffective power delivery and heat removal. This is important because high performance chips are projected to dissipate more than 100W/cm² and require more than 100A of supply current. Consequently, when such chips are stacked, the challenges in power delivery and cooling become greatly exacerbated. Thus, revolutionary interconnection and packaging technologies will be needed to address these limits [1].

This paper reports, for the first time, the configuration, fabrication, and experimental results of a 3D integration platform that can support the power delivery, signaling, and heat removal requirements for high-performance chips. The key behind this 3D platform is the ability to process integrate, at the wafer-level, electrical and microfluidic interconnection networks on the wafer containing the electrical circuitry and assemble such chips using conventional flip-chip technology.

Introduction

Three-dimensional (3D) die stacking increases transistor density and chip functionality by vertically integrating two or more dice. Furthermore, 3D integration also improves interconnect speed by decreasing interconnect wire length, enables smaller system form factor, and reduces power dissipation and crosstalk.

As transistor technology continues to scale and integration density increases, one of the main performance limiters of an IC chip will be heat removal. Because the reliability and performance of transistors and interconnects depend on operating temperature, the need to cool electronics and diminish hot spots has never been greater. With the continued scaling of device features and increased power density, chip cooling has become increasingly more difficult and costly.

The International Technology Roadmap for Semiconductors (ITRS) projects that power dissipation will reach 151 W and 198 W for cost performance and high-performance applications, respectively, by 2018 at the 18 nm technology node [2]. Consequently, it is not likely that conventional heat removal techniques will meet the power density, heat flux, and thermal resistance needs of future high performance microprocessors.

Liquid cooling using microchannels is promising to meet the thermal management requirements of future high-performance microprocessors due to its high heat transfer coefficient [3]. Although a number of researchers have explored the advantages of using liquid cooling to mitigate future thermal management problems [4-8], there are many unknowns for its implementation, especially for 3D integrated systems. Some of these unknowns include fabrication of an on-chip microfluidic heat sink and integration of electrical through-silicon vias (TSVs), where to place fluidic I/O interconnects for 3D chips, how to supply fluid to and extract fluid from microchannels embedded in a 3D stack, and how to assemble 3D ICs with microfluidic functionality.

This paper outlines process integration and assembly technologies for a proposed microfluidic liquid cooling configuration to cool three-dimensional ICs.

This paper is organized as follows: Section II of this paper presents a novel microfluidic cooling scheme for 3D ICs and outlines the fabrication process of the system components. In Section III, chip stacking, chip to substrate assembly, and fluidic sealing are discussed. Section IV outlines the testing of the system. Section V presents concluding remarks.
Section II: Novel Microfluidic Network Configuration for 3D ICs

A. 3D Microfluidic Network Cooling Scheme

We investigate, for the first time, microchannel heat sink integration into each stratum of a 3D stack to enable cooling of >100W/cm² of each high power density chip. Figure 2 shows a proposed microfluidic network cooling scheme that has the potential to be used for cooling three-dimensional ICs. Each silicon die of the 3D stack contains the following features: 1) a monolithically integrated microchannel heat sink, 2) through-silicon electrical (copper) vias (TSEV) and through-silicon fluidic (hollow) vias (TSFV), the latter used for fluidic routing in the 3D stack, and 3) solder bumps (electrical I/Os) and microscale polymer pipes (fluidic I/Os) on the side of the chip opposite to the microchannel heat sink [9]. Microscale fluidic interconnection between strata is enabled by through-wafer fluidic vias and polymer pipe I/O interconnects. The chips are designed such that when they are stacked, each chip makes electrical and fluidic interconnection to the dice above and below. Consequently, power delivery and signaling can be supported by the electrical interconnects (solder bumps and copper TSVs), and heat removal for each stratum can be supported by the fluidic I/Os and microchannel heat sinks.

Figure 2. Schematic of proposed chip-scale microchannel heat sink for 3D integrated circuits.

The following sections report the details of the 3D configuration and fabrication processes used to enable the wafer-scale batch fabrication of all interconnection modes described above.

B. Integrated Microchannel Heat Sink and Electrical TSVs

B. Dang et al [9-11] recently explored novel fabrication approaches for developing a CMOS-compatible, on-chip microchannel heat sink and thermo-fluidic I/O interconnection and packaging methods needed to implement chip-level microfluidic cooling. For 3D technology, electrical TSVs are needed for power delivery to chips in the 3D and to enable communication between chips in the 3D stack. The following schematic (Figure 3) outlines the process flow for integrating microchannels with electrical TSVs.

Figure 3. Schematic of wafer-level integration of microchannels, through-silicon fluidic vias, and electrical through-silicon vias to enable 3D system integration using liquid cooling. (a) Deposit oxide on front-side of wafer. (b) Pattern and etch through-silicon vias. (c) Cu seed layer evaporation and electroplating. (d) Fluidic TSVs and microchannel trenches are etched into the back side of the wafer. (e) Spin coat and polish Unity sacrificial polymer. (f) Spin coat and pattern Avatrel polymer sockets. (g) Simultaneous curing of Avatrel polymer and thermally decomposition of sacrificial polymer.

The process begins by depositing a 3µm thick layer of silicon-oxide on the front side of the wafer as a through-silicon via etch-stop layer (Figure 3a). TSVs are patterned and anisotropically etched into the back side of the silicon wafer in an inductive coupled plasma (ICP) etching tool (Figure 3b). After thermally growing a 1µm layer of oxide on the TSV sidewalls, a Ti/Cu seed layer is evaporated on the front-side of the wafer. After which, copper is electroplated in the TSVs (Figure 3c). Next, using two lithography steps, fluidic TSVs and microchannel trenches are etched into the back side of the wafer (Figure 3d). Subsequently, Unity sacrificial polymer (Promerus, LLC) is spin-coated on the wafer, filling the fluidic TSVs and microchannels. Afterwards, mechanical polishing is performed to planarize the surface (Figure 3e). Next, 15 µm of Avatrel 2090P polymer (Promerus, LLC) is spin-coated on the wafer, filling the fluidic TSVs and microchannels. Finally, the Avatrel polymer is cured, and the Unity sacrificial polymer is thermally decomposed simultaneously, making the process CMOS-compatible (Figure 3f). Figure 4 shows a cross-sectional optical image of a sample after the previously described processing steps are completed. The non-optimized microchannels are 200µm tall and 100µm wide (Figure 5), and the copper TSVs have a 50µm diameter. Platinum resistors are fabricated on the front-side of the wafer to facilitate heating and temperature sensing, as shown in Figure 6.
are stacked and a coolant is circulated from the top chip, through the 3D stack, and out of the bottom of the substrate.

Figure 7 shows a schematic of the fabrication process flow for a silicon die with integrated polymer sockets, through-wafer fluidic interconnects, thermofluidic I/O interconnects, and electrical I/O interconnects. The process begins by sputtering a 300/10000/300Å titanium/copper/titanium (Ti/Cu/Ti) metal layer, where Ti serves as an adhesion promoter between Cu and silicon (Figure 7a). The metal is patterned using a wet etch process (Figure 7b). Next, 1µm of oxide is deposited on the back side of the wafer as a polymer adhesion layer (Figure 7c), and 3µm of oxide is deposited on the front side as a through-silicon via etch-stop layer (Figure 7d). Next, 15µm of Avatrel 2090P polymer is spin coated onto the wafer (Figure 7e). Afterwards, polymer sockets are patterned on top of the metal (Figure 7f, Figure 8). The first layer of Ti is removed using a wet-etching process. Through-wafer fluidic vias are patterned and anisotropically etched into the back side of the silicon wafer in an ICP etching tool (Figure 7g, Figure 8); the etching stops at the etch-stop layer on the front side of the wafer. Next, a 12µm layer of Avatrel polymer is spin coated and patterned on the front side of the wafer and used as a passivation layer (Figure 7h). After sputtering a 300/2000/300Å Ti/Cu/Ti seed layer and electroplating a
2µm nickel under-bump metallurgy layer, 50 µm C4 solder bumps are electroplated for area-array electrical interconnects (Figure 7i, Figure 10). Afterwards, a 60 µm layer of Avatrel polymer is spin coated onto the front side of the wafer and used to pattern polymer pipes, which serve as thermofluidic I/O interconnects (Figure 7j, Figure 11). Finally, the oxide layer covering the through-wafer fluidic vias on the front side of the wafer is removed using a wet etch process to allow fluidic circulation.

Section III: Assembly of 3D Prototype Chips

A. Flip-chip Assembly Process

A challenge in such a 3D configuration is the flip-chip bonding process, especially since one must be able to provide fluidic sealing to prevent leakage. This section discusses the flip-chip die-to-die bonding processes that enable such integration of the components of the prototype 3D stack discussed in Section II.C. Fabrication and assembly of the prototype 3D chip stack allow the issues of 1) where to place fluidic I/O interconnects for 3D chips, 2) how to supply fluid to each die in the 3D stack, and 3) how to assemble 3D ICs with microfluidic functionality to be investigated.

The 3D prototype contains silicon dice with integrated through-wafer fluidic interconnects, thermofluidic polymer pipe I/O interconnects and high density solder bump electrical I/O interconnects on the front side of the dice, and polymer sockets on the back side of the dice, as shown in Figure 12. The silicon substrate contains copper pads, polymer sockets, and integrated fluidic TSVs.
two into contact with a compression force of 200g, and elevating the temperature of the chip and the substrate to 230°C and 150°C, respectively. The bonding process parameters are listed in Table 1. The fluidic I/Os and electrical I/Os are assembled simultaneously. The 250µm diameter polymer pipes are aligned to the 270µm diameter polymer sockets on the substrate, and the 50µm diameter solder bumps are aligned to the copper traces and 60µm polymer sockets on the substrate (Figure 12). In addition to serving as electrical and fluidic I/Os, the solder bumps and polymer pipes provide mechanical interconnection between the bottom die and the substrate and between the dice in the 3D stack.

Because copper traces and polymer sockets are fabricated onto the back side of the first die, the second die is assembled onto the back side of the first die using the same bonding recipe (temperature and force). The self-alignment property of solder increases the alignment accuracy of the die to the substrate. Patterned silicon dioxide on the substrate contains the solder during reflow. The process used for assembly of the 3D prototype is thus compatible with conventional flip-chip bonding. Figure 13 shows an SEM cross-sectional image of a 3D stack of two microfluidic chips. Figures 14a and 14b show infrared microscope images of through-silicon fluidic via alignment of the two chips. Figure 15 shows a 3D stack of two chips assembled to a substrate and a 3D stack of four chips assembled to a substrate.

<table>
<thead>
<tr>
<th>Assembly Parameters and Sequence of Steps</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-heating Temperature of Substrate</td>
<td>140°C</td>
</tr>
<tr>
<td>Pre-heating Temperature of Die 1</td>
<td>180°C</td>
</tr>
<tr>
<td>Compression Force</td>
<td>200g</td>
</tr>
<tr>
<td>Bonding Temperature of Substrate</td>
<td>150°C</td>
</tr>
<tr>
<td>Bonding Temperature of Die 1</td>
<td>230°C</td>
</tr>
<tr>
<td>Pre-heating Temperature of Substrate and Die 1</td>
<td>140°C</td>
</tr>
<tr>
<td>Pre-heating Temperature of Die 2</td>
<td>180°C</td>
</tr>
<tr>
<td>Compression Force</td>
<td>200g</td>
</tr>
<tr>
<td>Bonding Temperature of Substrate and Die 1</td>
<td>180°C</td>
</tr>
<tr>
<td>Bonding Temperature of Die 2</td>
<td>230°C</td>
</tr>
</tbody>
</table>

Table 1. Bonding Process Parameters

Figure 13. Cross-sectional SEM image of 3D microfluidic chip-to-chip bonding.

Figure 14a. Top-view IR-microscope image of 3D stack.

Figure 14b. Tilted IR-microscope image of 3D stack.
B. Fluidic Sealing

After assembly, to seal the 80 fluidic I/Os on the front side of each chip, an epoxy-based underfill is applied at the edges of chip (Figure 16). The underfill provides a stronger mechanical connection between each interface. For this application, most importantly, underfill is used for the purpose of sealing the fluidic interconnect interfaces between the die and the substrate and between the dice in the 3D stack.

Section IV: Testing

To test the reliability of the fluidic sealant, a mechanical pump and fluid inlet pipe can be attached to the back side of the top-most die in the 3D stack and used to pass fluid through the chip stack (Figure 17-18). This fluid delivery method is an alternative to the fluid delivery method described in Figure 2. Fluid can be delivered from the top chip, through the 3D stack, and out of the bottom of the substrate with no leakage at the chip-to-chip and chip-to-substrate interfaces.
Section V: Conclusions

Heat removal technologies are among the most critical needs for 3D integration of high-performance microprocessors. Liquid cooling represents a promising solution for meeting future thermal management requirements of high-performance 3D chip stacks. In this work, we outline the configuration, fabrication, and experimental results of a 3D integration platform that can support the power delivery, signaling, and heat removal requirements for high-performance chips. The core technologies behind this 3D platform, the ability to process integrate, at the wafer-level, electrical and microfluidic interconnects and assemble microfluidic chips using conventional flip-chip technology are demonstrated.

Acknowledgements

This work has been carried out as part of the Interconnect Focus Center Research Program at the Georgia Institute of Technology, and is supported in part by the Microelectronics Advanced Research Corporation (MARCO), its participating companies, and DARPA under contract 2003-IT-674, and in part is based upon work supported by the National Science Foundation under Grant No. 0701560.

References