Three-dimensional stacking IC packaging technology for NAND-flash memory

IOP Conf. Series: Materials Science and Engineering **537** (2019) 032064 doi:10.1088/1757-899X/537/3/032064

Three-dimensional stacking IC packaging technology for NAND-flash memory

V V Perminov[1], V V Putrolaynen[1], A S Shtykov[1] and A V Yartsev [1,2]

[1] Petrozavodsk State University, 31 Lenina str., Petrozavodsk 185910, Russia
[2] GS Nanotech, 11 Industrialnaya str., Gusev 238050, Russia

Email: teh-val@yandex.ru

Abstract. The article presents the technology of three-dimensional multi-chip packaging and testing of NAND memory module implemented at GS Nanotech in collaboration with Petrozavodsk State University. The main technological operations and quality management methods at each stage of memory modules manufacture are described. These memory modules can be used for solid-state drives production.

1. Introduction

As the storing data volume and requirements for high-speed data access increases, the information technology industry abandon the hard disk drives (HDD) in favor of NAND-flash memory-based solid-state drives (SSD). Various data density increasing technologies are applied to reduce the cost of data storing on SSD. Such technologies are multi-level cell (MLC) [1], three-dimensional flash memory architectures at die level (V-NAND, BiCS) [2], [3], and multi-chip packaging [4]. The first two approaches can only be implemented by flash memory manufactures, of which there are currently only four companies in the world (Samsung, Toshiba, SK Hynix, and Micron/Intel). Whereas for local manufacturing of the high data density flash memory module, the multi-chip packaging can be used. All these technologies can be successfully applied together, but high complexity of design and manufacturing of NAND-flash memory can lead to failures during operation under certain conditions [5]-[6]. To prevent such situations, it is necessary to perform testing of finished products before shipping them to the consumer, taking into account all requirements. This paper briefly describes the technology of three-dimensional multi-chip packaging and testing of NAND memory module implemented at GS Nanotech.

2. Manufacture

The first process is the pre-assembly which includes the operations of thinning NAND memory wafers down to less than 100 microns, attaching them to frames with a die attach film (DAF) and dividing them into separate dies. The thinning is carried out in free steps: rough grinding, fine grinding, and dry polishing. The used equipment automates the thickness control at each step, and the operations of unloading thick wafers from and loading thinned ones into the cassettes. The low die thickness requirement is caused by stacking up to eight dies into the module of about 1 millimeter thick. Such thin wafers are easy to bend and prone to break. So mounting them to frames with DAF is carried out by semi-automated setup (figure 1). The DAF consists of two layers: the adhesive layer, on which wafer is mounted and which is used to stack dies onto each other, and the base layer, which holds dies after

dividing. The dividing performs by sawing in two steps. At the first step the wafer is incised at 60–80% of the wafer thickness, and at the second step, the remaining part of the silicon and the DAF film are cut out by a thinner saw blade. After that, the DAF is irradiated by UV to decrease the adhesion between the base and adhesive layers for safety die separation from the frame. The used automatic dicing saw provides automatic dicing, control of the blade wear and frontside chipping measurement. The visual control of cut quality, chipping, cracks, and damaged dies is conducted by the operator using the optical microscope (figure 2). In our previous work, we propose the method of automation of the chipping inspection [7].

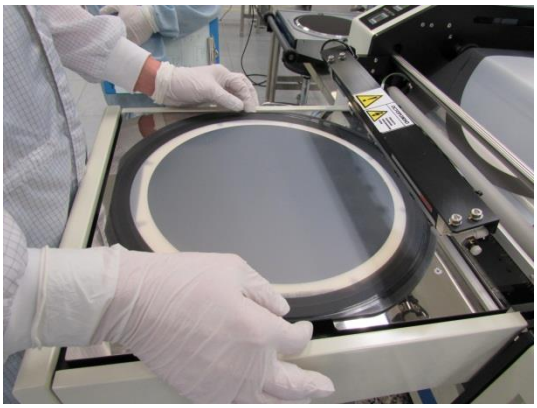

Figure 1. The operation of mounting thinned wafer to frame with DAF.

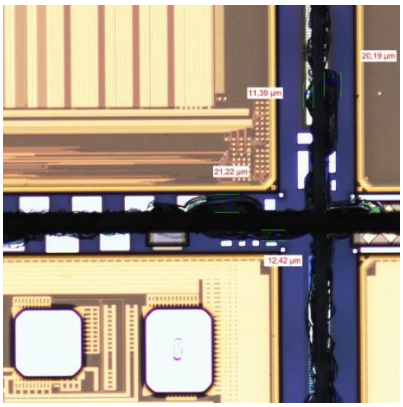

Figure 2. The example of the image of frontside chipping captured by an optical microscope.

The next stage is the attach print which consist in mounting dies to the substrate. To optimize the manufacturing process, the array of the substrates is united into the multiplied printed circuit board. The substrate has pads for electrical connection to the dies on the top side and pads for BGA solder balls on the bottom side. The dies with the adhesive layer of DAF are pushed up by the ejection system with needles and lifted from the base layer using a vacuum pick-up tool (figure 3). Then they are placed on the stack with an offset, in such a way that the contact pads of the underlying dies are open (figures 4, 5). All these operations are done by the automatic setup. The setup also provides automatic control of die placing accuracy in terms of the coordinate and angle using the computer vision methods.

After attaching the dies, the DAF film is polymerized in the oven. Compared with the use of a liquid adhesive, the using of the DAF reduces the risk of contamination of the dies contact pads in case of adhesive excess, or the formation of air voids under the dies, in case of its insufficiency. However, it is needed to thoroughly select the ejection system configuration, as the needles may pierce the DAF or break the die. Also, the choosing of a pick-up tool rubber is important, because it should provide enough force during die lifting and at the same time do not bend the thin die by vacuum, as it may lead to void formation under the attached die. In addition, as the DAF adhesive layer is not so fluid as a liquid adhesive, planarized substrates are needed to be used to avoid voids formation.

The electrical interconnections between dies and substrate are formed at the wire bonding process (figure 6). For this purpose, thermosonic ball-wedge and wedge-on-ball methods are applied to form die-to-substrate and die-to-die connections respectively. Before wire bonding, plasma cleaning of the substrate is done in order to remove mechanical, organic impurities and oxide films at the contact pads and activate the surface. Wire bonding is carried out as follows. The wire is threaded through capillary. At the wire end, the ball is formed by an electrical spark. After that, the ball is pressed to the die pad under application of heat and ultrasound. The “wedge” is formed at substrate’s pad or at the ball on the another die’s pad by the rim of the capillary under application of the same forces. As the height of the BGA package is limited at about 1 millimeter, and the stack of the dies and substrate occupies most of it, it is needed to form a low loop of wire with height less than 50 micrometers. In order to maintain high

bond quality and reliability, the process parameters are controlled by subjecting part of each lot to the wire pull test and the ball shear test.

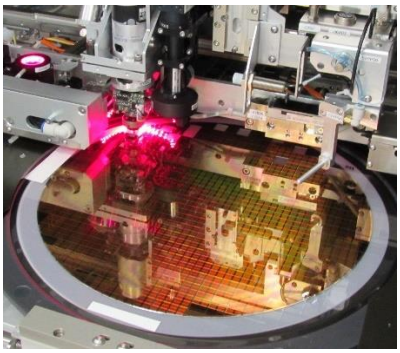

Figure 3. The picking-up of the die from the frame.

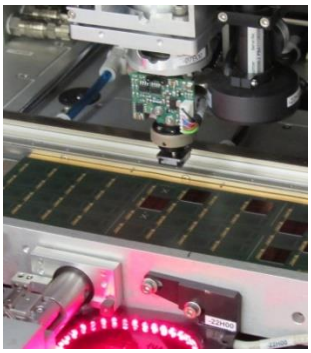

Figure 4. The stacking of the dies on the substrate.

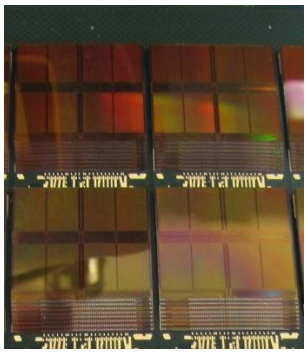

Figure 5. The stack of eight dies.

For protecting dies and interconnections from mechanical, chemical, electrical and other negative impacts of the environment, the molding process is carried out (figure 7). After clearing and activation of the surface by plasma, sealing by epoxy molding compound is performed. Then, the molded substrates are kept for several hours in the oven for compound polymerization, which increases the stability of its mechanical and electrical characteristics. At this stage following inspection methods are applied. The X-ray inspection system is used to detect breaks and jams of the wire interconnections. The acoustic microscope is applied to reveal the air voids and delamination. The warpage is measured by the non-contact optical profilometer.

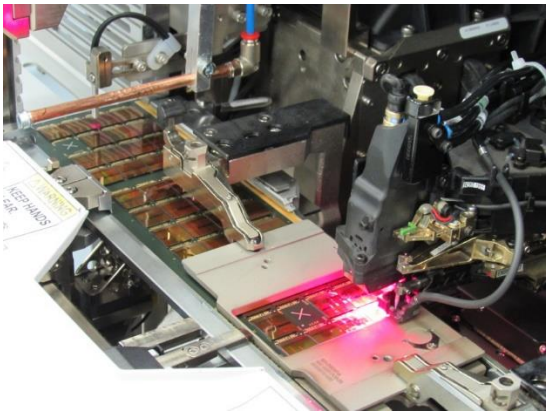

Figure 6. The wire bonding process.

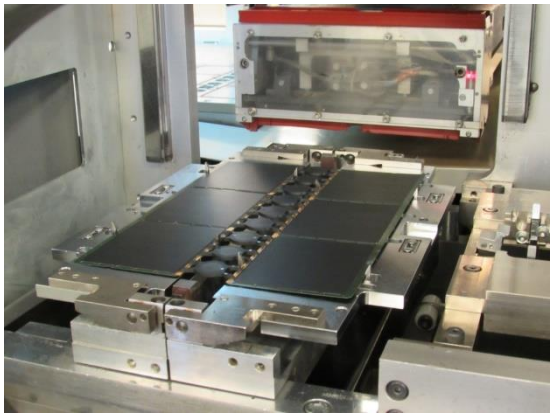

Figure 7. The molded substrates.

At the ball placing process, the solder balls are attached to the bottom side of the substrate. First, plasma cleaning is carried out. Using automatic setup the contact pads on the substrate's bottom is coated by flux, then solder balls are installed using a template with a vacuum pickup, and reflow of the balls is performed. By computer vision the setup checks for defects of balls placing before the reflow process. At the last stage, the singulation of the molded substrate into individual NAND memory BGA modules is performed.

3. Testing

The increasing complexity of the NAND-memory design and manufacture can lead to devices failures while operating under certain conditions. In order to prevent such situations, besides quality inspection at all stages of manufacture, it is necessary to perform functional testing of end products, before shipping them to the consumer. Although currently there are many commercial automatic test equipment solutions for NAND-memory modules, however, the using of that systems for testing small batches of

products (1000-10000 units) is not economically feasible. In this work, a modular system for testing multi-die NAND-based memory modules based on FPGA is used. Its architecture makes it possible to effectively scale the testing system (by now, simultaneous testing of up to 144 memory modules is carried out) and to flexible tuning its functionality. This system had been developed for the purpose of testing pilot batches of NAND memory modules, during the packaging technology processes development. The system executes testing programs, analyzes the results of the tests and generates a report for each memory module, while allowing both to create new testing programs and to store them for later use and analysis. The system is equipped with sockets, allowing the installation and removal of the BGA-memory modules without soldering and special tools usage. The more detailed description of the system was given in our previous work [8].

The testing of manufactured NAND memory modules is performed as follows. Firstly, the modules are tested by the system for open/short conductors, consumption currents at nominal voltage, leakage currents in the inputs/outputs circuits of the module to detect problems in the power supply and input/output circuits that occurred at the packaging process. Then, the functional verification test of fault-free modules is performed. It includes the program/read operation of specified data patterns and measurement of commands execution time. In the case of functional or time-out errors, the memory block is marked as bad-block. If the number of bad-blocks per LUN (Logical Unit Number) exceeds a certain value, the testing process is stopped, and the memory module is rejected.

Next, the memory modules that had passed the functional test are subjected to moisture sensitivity level (MSL) testing, thermal cycling (TC) testing and high-temperature storage life (HTSL) testing. The memory module lots are tested for MSL according to the JESD22-A113 standard (during 220 hours at 30 °C, 60% RH), using the humidity chamber. Then lots are divided into two groups. The first group is tested for TC in the thermal shock chamber in accordance with JESD22-A104 (400 cycles with a temperature change from -55 to +125 °C). The second one is tested for HTSL according to ESD22-A104 (1000 hours at 150 °C).

Finally, repeated functional testing is performed. The memory modules passed the final functional testing are marked as good.

The application software had been developed to implement the functional test procedure management. The software allows us to flexibly customize the testing process, such as data exchange and FPGA control, which provides high performance, functionality, and control of errors that occur.

4. Conclusion

This work presents the developed and proven technology of three-dimensional multi-chip packaging and testing of NAND memory module. The developed technology provides the manufacture of the high data density NAND-flash memory module with high reliability complying with international standards. These memory modules can be used for SATA, M.2, U.2, and other advanced solid-state drives production.

Acknowledgments

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (project no. 14.580.21.0009, unique identifier RFMEFI58017X0009).

References

[1] Liu S and Zou X 2017 QLC NAND study and enhanced Gray coding methods for sixteen-level-based program algorithms *Microelectronics Journal* **66** 58–66
[2] Yamashita R et al. 2017 *2017 IEEE Int. Solid-State Circuits Conf. (ISSCC)* (USA: Penmor Lithographers) pp 196–7
[3] Im J-W et al. 2015 *2015 IEEE Int. Solid-State Circuits Conf.-(ISSCC) Digest of Technical Papers* (USA: Penmor Lithographers) p 130
[4] Huang H and Micheloni R 2016 *3D Flash Memories* ed R Micheloni (Dordrecht: Springer Netherlands) pp 261–79

[5] Wang Y, Liu Y, Li M, Tu K N and Xu L 2017 *3D Microelectronic Packaging: From Fundamentals to Applications* ed Y Li and D Goyal (Cham: Springer International Publishing) pp 375–420
[6] Tu K N 2011 Reliability challenges in 3D IC packaging technology *Microelectronics Reliability* **51** 517–23
[7] Perminov V, Putrolaynen V, Belyaev M, Pasko E and Balashkov K 2018 Automated image analysis for evaluation of wafer backside chipping *The International Journal of Advanced Manufacturing Technology* **99** 2015–23
[8] Podryadchikov S, Putrolaynen V, Belyaev M, Chuvstvin M and Tabachnik I 2019 FPGA-based testing system of NAND-memory multi-chip modules *Microelectronics Journal* **83** 73–6