

A 1.2 V 20 nm 307 GB/s HBM DRAM With At-Speed Wafer-Level IO Test Scheme and Adaptive Refresh Considering Temperature Distribution

Kyomin Sohn, Won-Joo Yun, Reum Oh, Chi-Sung Oh, Seong-Young Seo, Min-Sang Park, Dong-Hak Shin, Won-Chang Jung, Sang-Hoon Shin, Je-Min Ryu, Hye-Seung Yu, Jae-Hun Jung, Hyunui Lee, Seok-Yong Kang, Young-Soo Sohn, Jung-Hwan Choi, Yong-Cheol Bae, Seong-Jin Jang, and Gyoyoung Jin

Abstract—A 1.2 V 20 nm 307 GB/s high-bandwidth memory (HBM) DRAM is presented to satisfy a high-bandwidth requirement of high-performance computing application. The HBM is composed of buffer die and multiple core dies, and each core die has 8 Gb DRAM cell array with additional 1 Gb ECC array. At-speed wafer level, a u-bump IO test scheme and an adaptive refresh scheme considering temperature distribution are proposed to guarantee test coverage and stable operation in a power-efficient manner.

Index Terms—Adaptive refresh, at-speed wafer level test, high-bandwidth memory (HBM) DRAM, microbump IO, TSV.

I. INTRODUCTION

FIG. 1 shows the trends of data rate per pin and data bandwidth per device of recently published DRAMs [1]–[3]. Demand for higher bandwidth DRAM continues to increase, especially in high-performance computing and graphics applications. However, it is difficult to increase pin data rate more than 10 Gb/s for the higher bandwidth. Multiple GDDR DRAMs could be a solution for higher bandwidth. In that case, high power consumption and routing congestion on PCBs become a big concern. Another way to increase bandwidth is to use many IOs, such as wide IO or high-bandwidth memory (HBM) [3], [4]. The number of HBM IOs (1024 IO) is far more than other conventional DRAMs (around 32 IO). Wide IO DRAM has also many IOs (512 IO), but it has a limitation in increasing memory capacity, because wide IO DRAM is stacked over processor. In addition, DRAM cell retention feature is very sensitive to temperature, so that it is not proper to be implemented over hot processor, such as GPU or high-performance processor.

In order to overcome these limitations, the HBM DRAM was recently introduced. HBM-DRAM uses TSV and interposer technologies enabling multiple chip stacks and wide

Manuscript received May 2, 2016; revised June 20, 2016; accepted August 6, 2016. Date of publication September 13, 2016; date of current version January 4, 2017. This paper was approved by Guest Editor Atsushi Kawasumi.

The authors are with Samsung Electronics, Hwaseong 445-701, South Korea (e-mail: kyomin.sohn@samsung.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2016.2602221

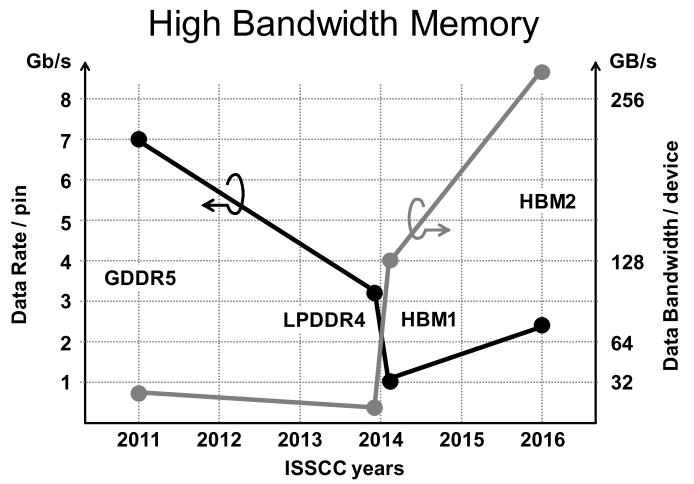


Fig. 1. Bandwidth trend of DRAM memory.

I/Os between the processor and memory: providing high capacity, low power, and high bandwidth. This paper proposes the second generation HBM to double the bandwidth from 128 to more than 256 GB/s and support pseudochannel mode and 8H stacks [5].

Table I shows the comparison of GDDR, HBM Gen1, and Gen2. The pin data rate of HBM is lower than that of GDDR. They are only 1 and 2 Gb/s. However, HBM has eight channels and 128 IOs per each channel. Totally, there are 1024 IOs per one HBM device. By this way, the bandwidth of HBM Gen2 is 256 GB/s per device, which is the double of HBM Gen1 and the highest bandwidth of DRAM history. Supply voltages of HBM are 1.2, 1.2, and 2.5 V for V_{DD_CORE} , V_{DD_IO} , and V_{PP} , respectively, which is the same as DDR4 SDRAM and lower voltages than GDDR5 for low power consumption. An interface method is unterminated CMOS swing, which is the same as LPDDR3. It comes from small pin capacitance by using short interposer channel and point-to-point configuration between HBM and processor. Bank composition of one channel is the same as that of GDDR5, which has 4 bank groups and 4 banks per each group, totally 16 banks.

HBM Gen2 has new functionalities in comparison with HBM Gen1. They are pseudochannel mode, 2H and 8H

TABLE I
COMPARISON TABLE OF GDDR5 AND HBM GEN1 AND 2

Items	GDDR5	HBM Gen1	HBM Gen2
Pin Data Rate	4~8Gbps	1Gbps	2Gbps
# of IO and CH	1CH, 32 IO	8CH, 128IO/CH	16pCH, 64IO/pCH
Bandwidth	16~32GB/s	128GB/s	256GB/s
Voltage (VDDC/VDDQ/VPPE)	1.35V~1.5V	1.2V/1.2V/2.5V	←
Interface	POD (VDDQ Term.)	CMOS (Un-terminated)	←
Banks	4banks/BG, 4BGs	←	←
Implemented new functions in this work			Pseudo channel, 2H/4H/8H, ECC storage, Implicit pre-charge, Lane remapping, ...

configurations, ECC storage, implicit precharge, lane remapping, and so on. In the pseudochannel mode, a legacy channel is divided into two pseudochannels, and the two pseudochannels share the command-address pins. Thus, one HBM has 16 pseudochannels instead of 8 legacy channels. To support various stack configurations, including 8H stacks, a new architecture is adopted for flexible density ranging from 16 to 64 Gb maintaining the same bandwidth. HBM Gen2 has optional ECC storage for storing parity bits, which is used for ECC calculation. Implicit precharge is for efficient utilization of command-address bandwidth. Lane remapping is to replace fault lane with redundant lane after assembly for high yield.

In this paper, we present HBM Gen2 DRAM in a 20 nm CMOS technology. As a measurement result, it shows 2.4 Gb/s/pin data rate in 1.1 V supply voltage, which means 307 GB/s bandwidth. This paper is organized as follows. Section II provides the chip architecture of stack, buffer, and core dies. Section III describes the thermal solution and TSV management technique. Wafer-level test methodologies, including direct access (DA) mode and IO design for excellence (DFx), are explained in Section IV. Section V shows a fabricated chip photo and measurement result. Section VI summarizes this paper.

II. ARCHITECTURE

A. Stacked DRAM

HBM is composed of stacked DRAM dies over a buffer die, as shown in Fig. 2. System-in-package (SiP) using HBM has typically one processor and two or four HBMs. One HBM has thousands of signals to communicate with processor, and channel length is recommended under 6 mm for low power and good signal integrity. Those requirements make the implementation of more HBMs than four difficult. HBM's power balls are directly connected through interposer and PCB to the bottom side of SiP. The proposed HBM Gen2 can have three types of stack configurations, 2H, 4H, and 8H. One core die has 8 Gb cells with 1 Gb cells of ECC storage. Therefore, 2H means 16 Gb, and 8H 64 Gb. However, the total height

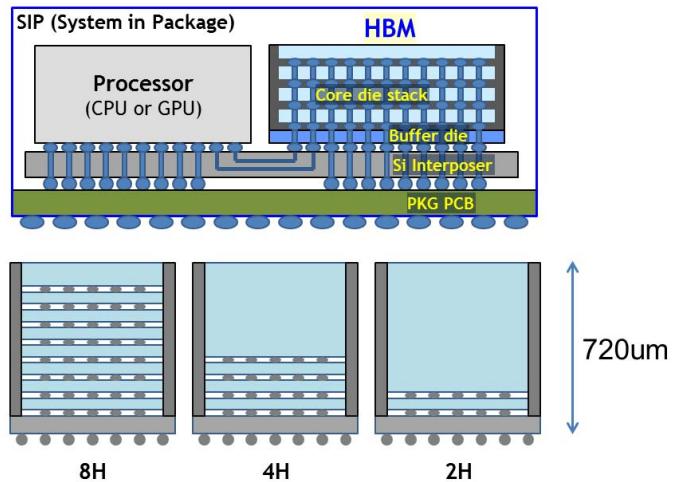


Fig. 2. SiP using HBM and 2H/4H/8H intersection.

of the HBM is the same regardless of the number of stacks for compatibility. The height of HBM is 720 μ m and it is necessary to apply thermal solution with processor, such as heat spreader or active cooler in all configurations. For this purpose, the thickness of the top core die is different depending on the number of stacks.

B. Buffer Die Architecture

The purpose of the buffer die is to: 1) provide routes from the TSVs related to the DRAM dies to the microbump IOs in the PHY area and 2) provide test functionality to system makers and DRAM vendors. DRAM tests are all executed through DA PADs in DRAM vendor site. Unlike other DRAMs, test feature is very important in HBM, because the microbump array cannot be directly probed. The test equipment to test microbump path directly is being developed, but currently it is not available. Fig. 3 shows that the test block covers all of the normal paths from the PHY to the TSV for each test item. Functionality and timing margins can be guaranteed by this architecture, since a DFT and a SerDes module with a phase locked loop (PLL) block are implemented for

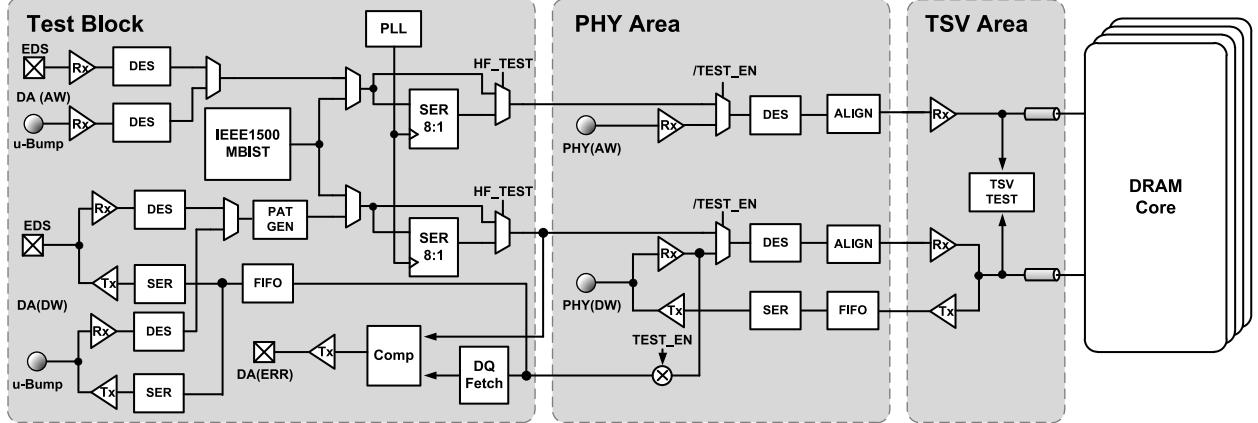


Fig. 3. Block diagram of buffer die's signal path.

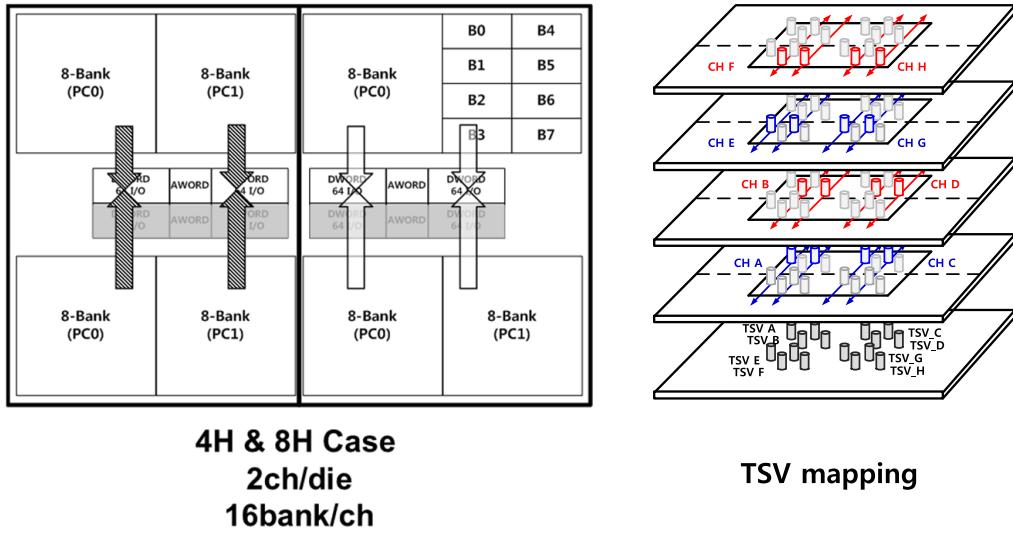


Fig. 4. Core die architecture of 4H and 8H configurations.

low-frequency (LF) test equipment. DA mode will be explained in more detail in Section IV.

BIST and IEEE1500 blocks support HBM tests on SiP, since it is difficult and inconvenient to test HBM after the chip on wafer (CoW) process is completed and the HBM is connected to a processor. System makers can test HBMs using these functions and isolate failure points when they occur.

C. Core Die Architecture

Fig. 4 shows the core die architecture of 4H and 8H configurations. Each die has a 9 Gb cell array, including 1 Gb cells for optional ECC. 4H means that four core dies are stacked over the buffer die. In this configuration, the HBM is composed of two channels and each channel has two pseudochannels (PC0/PC1), which consists of 16 banks (4 bank groups and 4 banks per group). In the middle area, there are TSV arrays to interconnect core die and buffer die. Depending on the number of floor, each core die uses different TSVs. The core die of the first and third floor uses upper TSV array and other dies use lower TSV array.

Fig. 5 shows the core die architecture of 2H configuration. Naturally, 2H case can use only four channels

(or eight pseudochannels) if the same configuration is applied as 4H and 8H cases. To meet the higher bandwidth requirements even in 2H case, additional circuitry for data management is introduced. One pseudochannel is divided into two different channels and each channel has eight banks, which is necessary to keep the same bandwidth as in the 4H/8H case. Unlike 4H and 8H cases, all TSV arrays are used in the 2H case. Upper channel uses the upper TSV array and lower channel the lower TSV array. There could be the performance degradation from 8 banks instead of 16 banks in the 2H case, but the bandwidth is more important for system performance.

This architecture supports 32 or 64 B prefetch depending on operating modes. In legacy mode, 32 B ($=128 \text{ I/O} \times 2$) for BL2 mode and 64 B ($=128 \text{ I/O} \times 4$) for BL4 mode are supported. Unlike legacy mode, pseudomode only supports 32 B ($=64 \text{ I/O} \times 4$) prefetch.

III. KEY SCHEMES FOR STACKED DRAM

A. Thermal Solution

HBM DRAM shows more than 40% power reduction compared with GDDR5 because of low supply voltage and lower power interface, which uses 1.2 V and untermination.

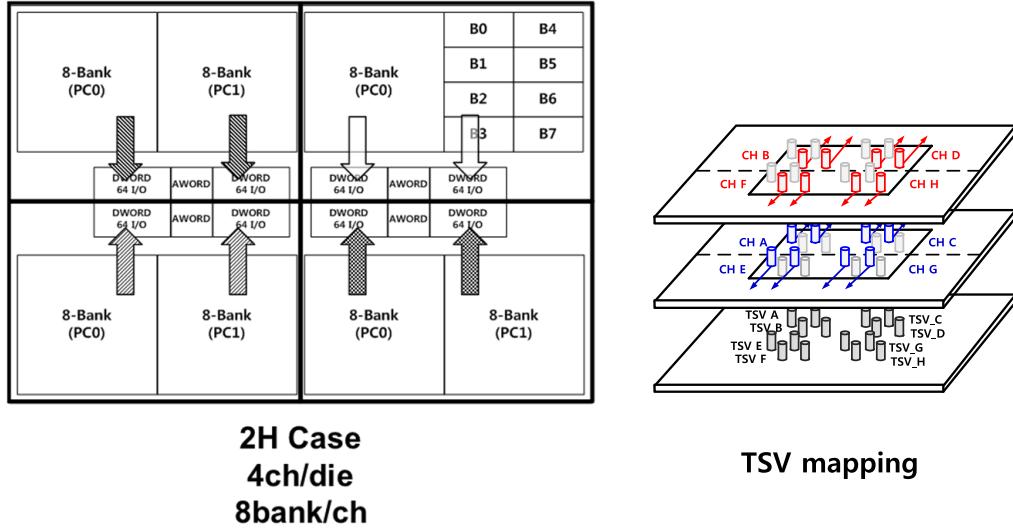


Fig. 5. Core die architecture of 2H configurations with full bandwidth.

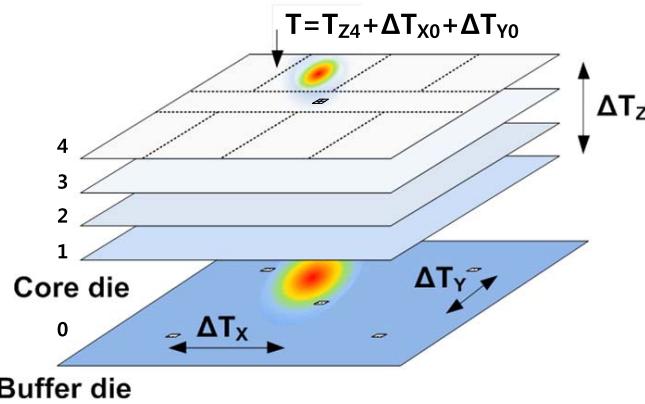


Fig. 6. Temperature distribution of HBM DRAM.

However, the power density is around three times worse than GDDR5 because of its small form factor. Since high power density causes thermal issues, which have critical impact on DRAM cell retention, both DRAM and controller should know the exact temperature condition of the DRAM cells so as to send proper refresh commands. Fig. 6 shows the temperature distribution inside HBM. There are two kinds of distribution, which are vertical and horizontal distribution. In a buffer die, there are five temperature sensors to detect the horizontal 2-D temperature distribution. In addition, each core die has a temperature sensor to detect the vertical distribution. It is difficult to implement multiple temperature sensors on the proper location of core die because of large cell array, which occupies most of chip. In addition, cooling solutions, such as heat spreader or liquid cooler, are implemented on top of HBM. Therefore, buffer die is the hottest die among dies of HBM in most cases. For example of the fourth stacked core die, as shown in Fig. 6, its temperature is determined by summing the temperature of the stack (T_{z4}) and the horizontal offset ($\Delta T_{x0} + \Delta T_{y0}$) measured from five different temperature sensors at buffer die.

To detect the temperature distribution of HBM and apply the proper refresh rate, the adaptive refresh considering

temperature distribution (ART) scheme is proposed, as shown in Fig. 7. The refresh controller for each die detects the spatial temperature difference (ΔT_X and ΔT_Y) and sets the proper refresh rate accordingly for the eight sections of the core die. The DRAM also sends the temperature code to the DRAM controller for setting the external refresh rate. By combining the external and internal refresh rate, the final effective refresh rate of the DRAM macro is determined as shown in the left side of Fig. 7. Thus, macros in thermal hotspots would have a high refresh rate for data retention, while refresh frequency in the cold macros is reduced. In other words, final refresh is determined by external refresh with adding or subtracting internal refresh, which enables fine control based on temperature according to locations in HBM. For banks on the stack with lower temperature than outputted max temperature code, external refresh commands are issued more than needed. In this case, according to internal refresh, which can skip some of external refresh, final refresh is determined for those banks. On the contrary, if internal temperature went higher before updating external refresh, internal refresh can perform more frequently than external refresh. By this way, the overall power consumption of DRAM refresh is reduced, and the DRAM cell data retention can be secured.

B. TSV Management

The bandwidth of HBM has been doubled over the previous generation HBM, yet the internal signal speed of TSV is limited due to the capacitance of TSV considering 8H case. To overcome the speed limit, HBM uses multiple TSVs in parallel, currently requiring more than 5000 TSVs, including power TSVs. Therefore, TSV defect detect and repair schemes are essential for massive TSVs; hence, the number of robust TSVs for test and repair must also increase [6]. A conventional robust TSV consists of three TSVs that majority vote on the actual output values. To reduce the number of TSVs required for the robust TSVs, a built-in self-test and selector scheme is proposed. Fig. 8 shows the signal waveform of this scheme, and Fig. 9 shows the diagram of the proposed robust TSV

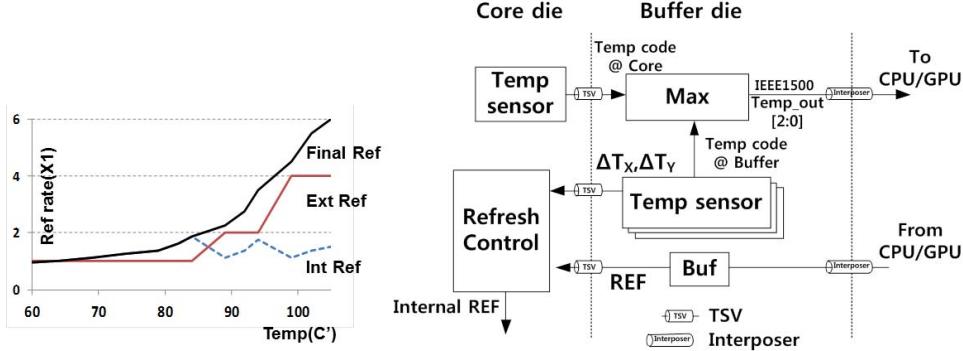


Fig. 7. ART scheme.

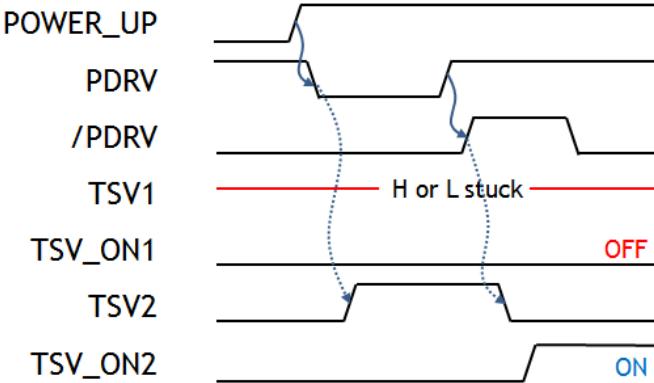


Fig. 8. Waveform of TSV defect autodetect and selection scheme.

module, including the self-test and selector logic. During the power-up sequence, test pulses are generated and transferred through TSVs at the buffer die first. Then, the output value of a flip-flop, which enables the receiver of TSV, is changed to the high level. If the output of TSV does not toggle, due to TSV failures, then the output value of flip-flop does not change and the receiver of the TSV under test maintains its OFF-state. The first TSV is selected automatically when two TSVs are all good. In other words, if TSV_ON1 is activated to high state as a result of the proposed scheme, TSV_ON2 is deactivated regardless of the defectiveness of TSV2 by additional XOR logic, which is not described in Fig. 9. The side effect of this scheme is that two TSVs for one signal are necessary. This scheme is adopted for the critical signal, such as RESET and several test signals. Conventional repair scheme is also used for secure TSV interconnection.

IV. WAFER-LEVEL TEST

Unlike conventional DRAM in package business, HBM uses a micropillar grid array (MPGA) and it is being tested as CoW before being individually separated. Before CoW stage, each core die and buffer die are tested as its own wafer for assembly to stacked die. A core die is tested mainly for cell access through electrical die sorting (EDS) pads. However, buffer die of HBM has many new functions to test, in which EDS pads are the only access at wafer level test. Therefore, DA microbump interface, which enables test even after CoW assembly, is replicated on EDS pad so that all the tests via DA path are capable on wafer.

A. Direct Access Mode

As shown in Fig. 3, a buffer die has microbumps in PHY area and DA microbumps in DA area as well as EDS pads for wafer test. The DA microbumps and the EDS pads have the same function; however, the operation speed is different due to the operation environment. The DA microbumps can cover up to at-speed operation, because they should be used for failure analysis in SiP with the same IO structure as PHY IO. But the EDS pads can operate at only LF, because the EDS pads are controlled by wafer test machine.

In DA mode, the HBM can operate as normal function mode through DA microbumps. The data path from PHY IO circuit to TSV is verified by read/write operation when a buffer die wafer is tested whether it is good enough to stack core dies on it. After stacked, DA test mode can determine the stack as a known good stack die (KGSD) by accessing cell and core on multiple core dies. In addition, failures after SiP assembly can be figured out by DA from DA balls routed to external PCB. To perform this function, the DA path (from DA microbumps to TSVs) is merged with microbump interface in PHY area to verify its own read/write path as well as PHY interfaces' read/write path. However, as described earlier, the EDS pads can only operate at LF due to test equipment. This problem is solved by introducing a SerDes scheme and PLL to multiply LF sources. During LF test modes, the DA path operation (pass/fail) is verified by read data operation after write date operation. On the other hand, high-frequency (HF) test via EDS pads generates data and clock up to at-frequency in buffer die by SerDes and PLL. However, the (wafer level) test machine cannot support such kind of HF. Therefore, data compare logic is implemented instead of read/write data via EDS pads, which compares the data pattern from SerDes and read-out data pattern after write operation. After that, pass/fail information is provided as a flag. By this operation, HF operation of DA path can be verified. For widening coverage of DA test modes, DA path contains all the buffer die's normal paths. An MUX is inserted between input buffer and deserializer to use data path after input buffer for write operation. In addition, for read operation, the output data are utilized, which comes from DQ driver's output to input buffer through noncontact microbump. With these operations, the read/write operations of PHY area at LF and HF (at-speed) can be verified. The DA test mode supports not only one-channel operation but also eight-channel operation.

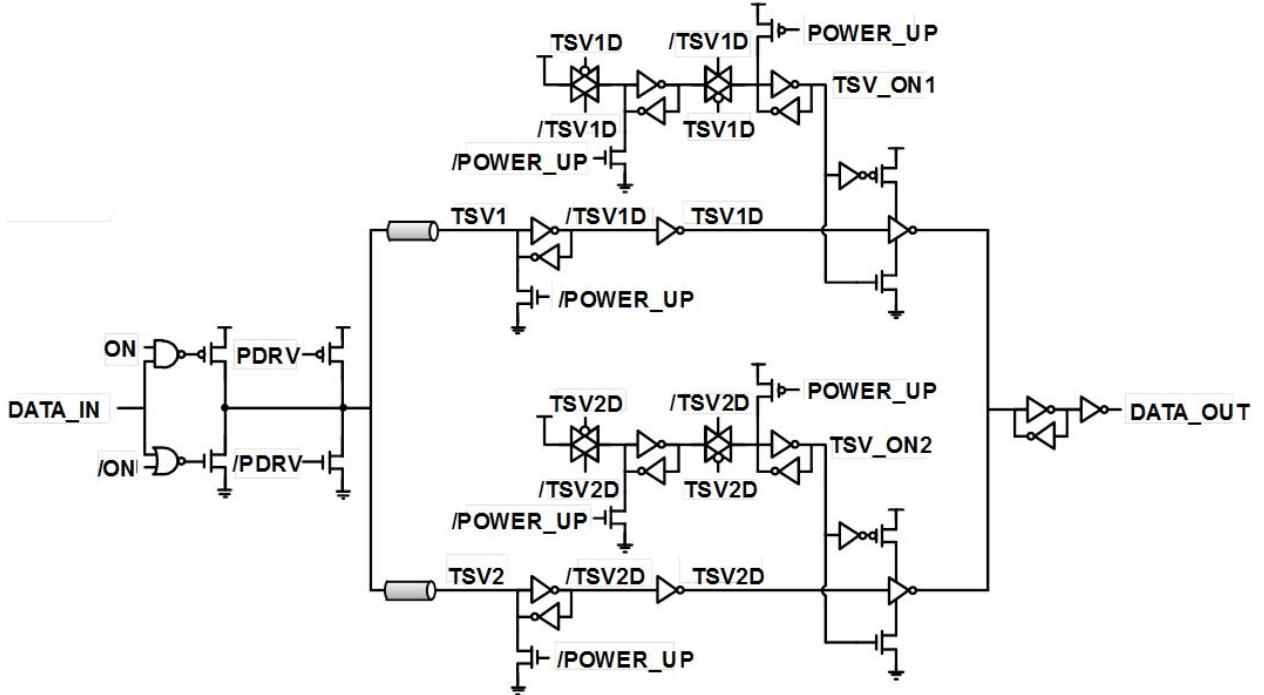


Fig. 9. Key circuitry of TSV defect autodetect and selection scheme.

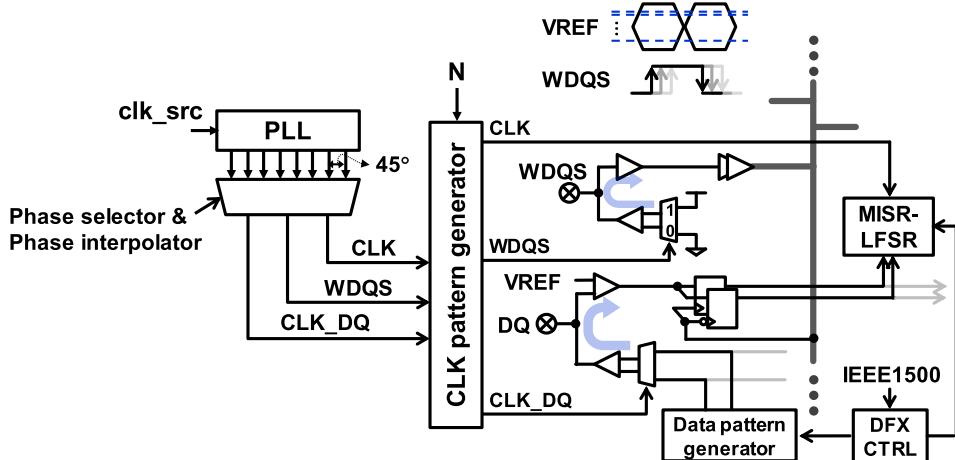


Fig. 10. Block diagram of IO DFx for DWORD.

B. PHY IO Design for Excellence

The HBM is assembled in 2.5-D SiP structure with host. Each of HBM, host, and interposer is made of different processes. Furthermore, SiP is another process, which was not encountered for each device during its manufacture. Therefore, the HBM has to be verified as KGSD when it is delivered to vendors. Since the HBM is composed of fine pitch microbumps for PHY interface, it is not possible to probe directly. In addition, the number of microbumps exceeds 1000, and it is difficult to find test socket, which supports the HBM microbump structure. Therefore, the HBM functions are verified in wafer level. As explained in Section IV-A, some parts of whole read/write paths are checked with their functionalities by DA test mode through pad loopback [7]. However, it is not enough for being granted as KGSD

in SiP, because the first test would be AC link margin test at operating speed, which is not covered by DA test mode. The AC link test includes loopback functions, which are multiple input signature register (MISR) write, linear feedback shift register (LFSR) compare, and LFSR read. Those functions are assumed that all I/O interfaces are PHY microbump interfaces, neither DA nor any other PADs.

Therefore, the IO DFx test scheme is proposed to test IO loopback with MISR write at-speed operation. To replicate microbump interface environments, DQS and DQ use the same driver structure, as shown in Fig. 10. From the state of microbump, the timings of DQ and DQS are conceptually identical with external signals. After buffer from DQ and DQS, normal write paths with MISR register are used to check AC margin. Data pattern can only be BL2 data in one test, such as "00", "01", "10", and "11". With this simple data

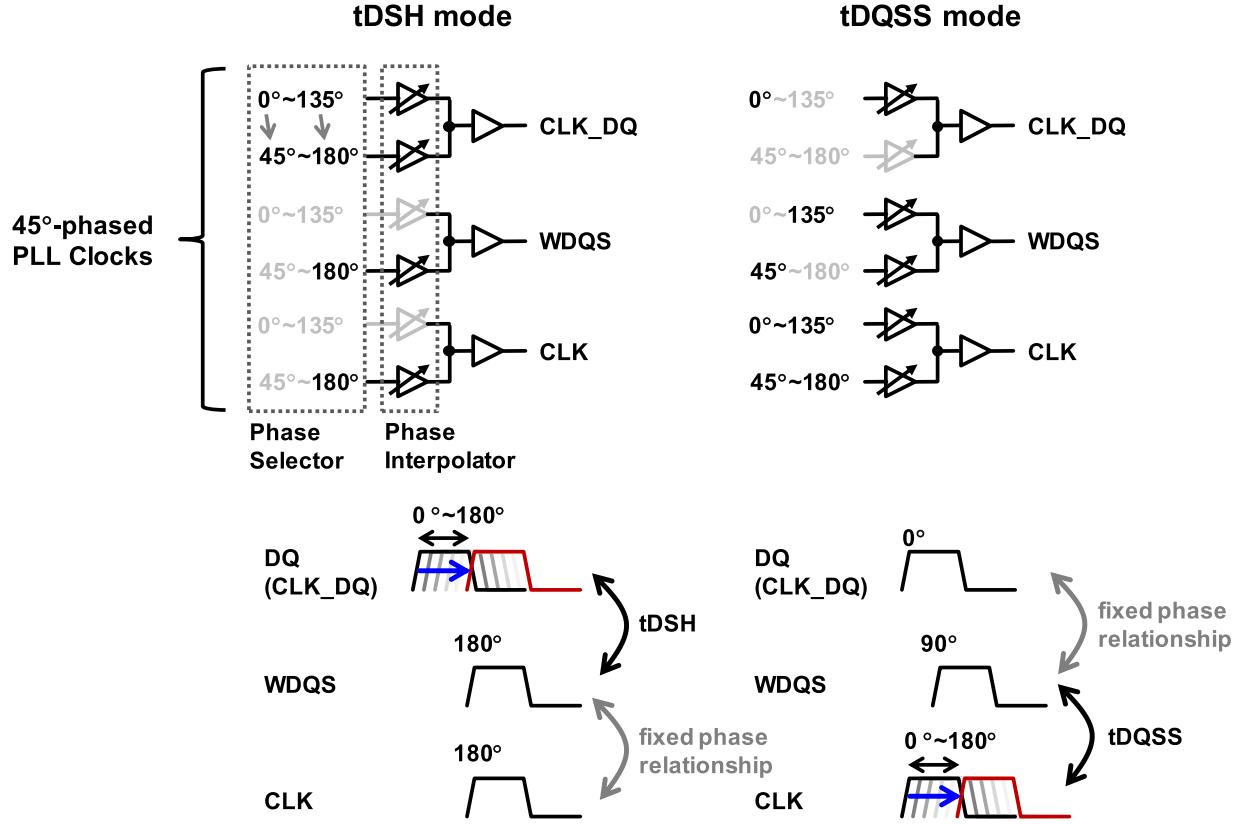


Fig. 11. IO DFx test mode: tDSH measure and tDQSS measure mode.

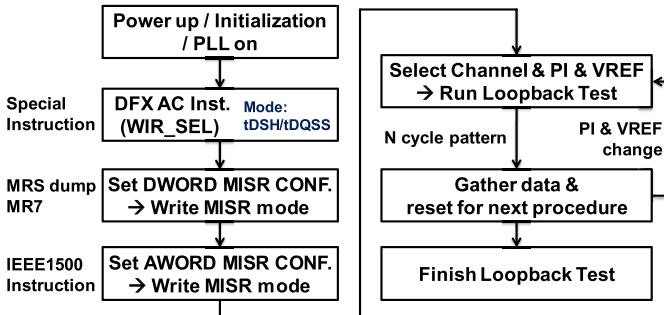
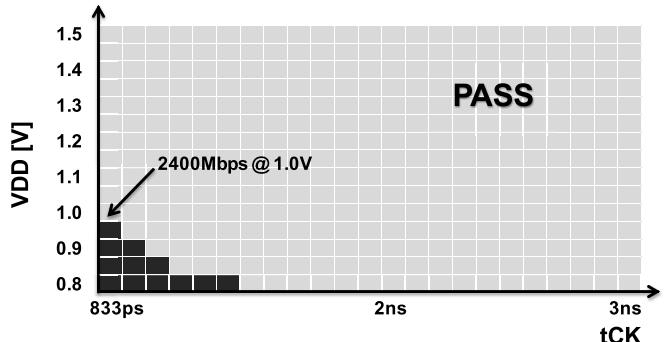
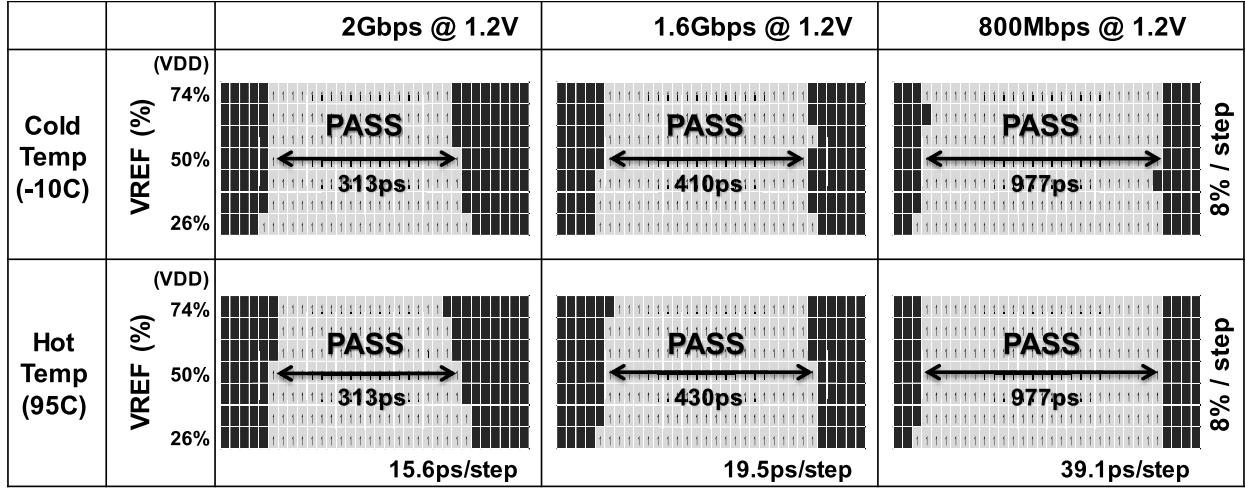
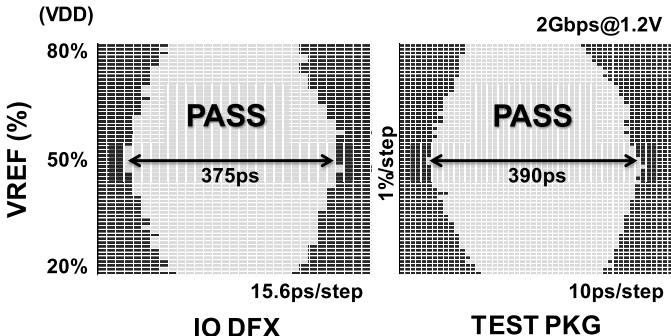


Fig. 12. Test sequence of IO DFx test.

stream, MISR write functions are performed to compress and generate signature. These signatures are captured and shifted out through IEEE1500 interface for checking its pass and fail. For write function, at least three different clocks are necessary, which are CLK for AWORD as well as the substitute of write command, WDQS for sampling inputted data, and CLK_DQ for data generation having timing relationships with DQS. An integrated PLL derives eight 45°-phased clocks for phase selector (PS) and phase interpolator (PI). In the PS and PI, three phase-movable clocks are selected for cases to test, which are setup/hold and DQSS mode. As shown in Fig. 11, in the tDSH mode, WDQS and CLK are fixed to 180°-phase and CLK_DQ for data timing generation can be freely movable by settings made by DFx control through IEEE1500 instructions. CLK_DQ phase is moving from 0° toward 180°. For instance

Fig. 13. Measured shmoos plot of tCK versus V_{DD} showing 2.4 Gb/s operation at 1 V.

of phases between 0° and 45°, two clocks of 0° and 45° are selected in the PS and the PI generates phase-shifted clock between them. Likewise, four sets of two 45°-phased clocks are used for evenly phased clock to test tDSH, which are 0°–45°, 45°–90°, 90°–135°, and 135°–180°. For each timing set, MISR signature is stored to indicate pass/fail. In the tDQSS mode, DQ and DQS timings are 90°-phased for maximum S/H condition, and CLK and DQS timing relationships are tested. CLK_DQ is fixed to 0°, WDQS is fixed to 90°, and CLK is moving from 0 to 180° for checking tDQSS. All these test sequences and setting are done by IEEE1500 special instruction. The test sequences are shown in Fig. 12, and this DFx AC test can check AC margin of DWORD as

Fig. 14. V_{REF} shmoo plots of IO DFX test at $-10^{\circ}\text{C}/95^{\circ}\text{C}$.Fig. 15. V_{REF} shmoo comparison between IO DFX test and ATE-based test.

well as AWORD. In HBM spec, internal V_{REF} is defined with eight cases, which are 38%–66% with 4% step. However, for better testability, this V_{REF} range is extended and resolution becomes finer while IO DFX test is enabled. As results, 2-D shmoos can be obtained by varying the programmable clock phases and V_{REF} from 20% to 80% of V_{DD} according to test loop sequences.

V. MEASUREMENT RESULTS

Fig. 13 shows the shmoos plots for tCK and V_{DD} , which results from IO DFX test. It achieves 2.4 Gb/s/pin operation at a 1 V supply voltage, which results in 307 GB/s of total bandwidth and it makes 1.2 TB/s in SiP with four HBM configurations. In Fig. 14, the IO DFX test results of V_{REF} shmoos are shown according to various temperatures and frequencies. The x-axis of the shmoos indicates 32 steps of PI attached to PLL, and the y-axis does V_{REF} levels proportional to external V_{DD} supply voltage, in this case V_{DDQ} . In this measured shmoos plot, one step of the y-axis is 8% of V_{DDQ} . The comparison between measured V_{REF} shmoos from IO DFX test and the one from ATE-based measurements via test package are shown in Fig. 15. Since the output terminal of MPGA is a microbump, it cannot be tested by a conventional test socket or board. In order to measure through

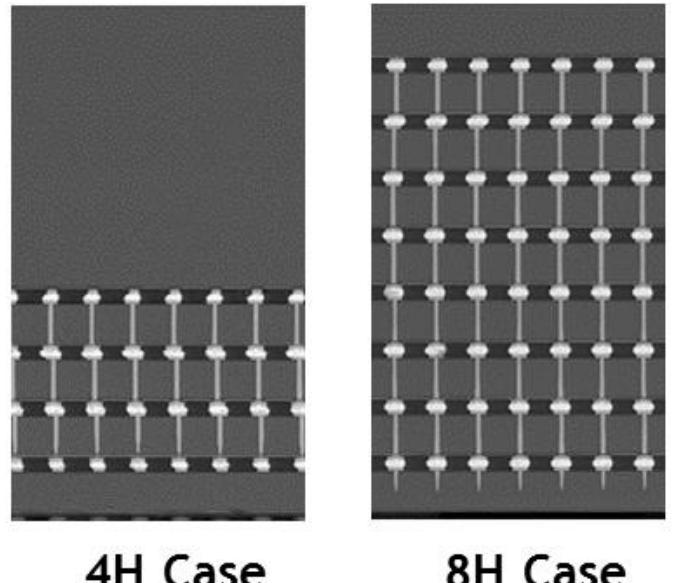


Fig. 16. Photos of cross sectional of 4H and 8H stacked dies.

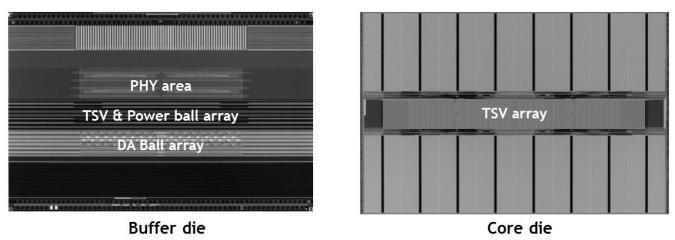


Fig. 17. Chip microphotographs of the core die and buffer die.

PHY interface, that is microbump, a test package using a silicon interposer die and HBM cube was used. Based on measurement results, IO DFX AC test results seem to have similar results with ATE-based test results.

The HBM chip is fabricated in a 20 nm DRAM process, and its size is $12 \times 8 \text{ mm}^2$. Supply voltages for $V_{DD}/V_{DDQ}/V_{PPE}$

TABLE II
PERFORMANCE SUMMARY

Item	Description
Process	20nm DRAM Process
Chip size	12mm x 8mm
Organization	Pseudo 16 channel 16 banks/channel 64 IOs/channel
Density	8Gb + 1Gb / chip (1Gb – ECC block)
Bandwidth	307GB/s/cube
Data Rate	2.4Gb/s/pin
Supply voltage (VDD/VDDQ/VPPE)	1.2V/1.2V/2.5V
Refresh rate	8k/32ms (Temperature Controlled)

are 1.2/1.2/2.5 V, respectively. VPPE is the wordline driver supply voltage. The refresh rate is 8 k/32 ms at room temperature. Fig. 16 shows the cross-sectional photos of 4H and 8H stacked dies. In addition, the chip microphotographs of the core die and buffer die are shown in Fig. 17. The performance of this HBM is summarized in Table II.

VI. CONCLUSION

In this paper, the second generation HBM was presented. The proposed 64 Gb HBM with 8H stack operates at 307 GB/s at 1 V. As results, with four HBM memories on SiP, total bandwidth of more than 1 TB/s can be achievable. By using ART scheme, cell refresh performance degradation caused by high power density was removed. Comparing with GDDR5, therefore, HBM could have twice the power efficiency from 6.6 to 12 W in case of 256 GB/s bandwidth. For mass production with 5k TSVs, a built-in self-test and selector logic is introduced. To verify bare-die wafer or stacked die wafer-level operations, DA test paths, including normal write/read path and IO DFX test scheme with at-speed operation, are proposed and implemented.

REFERENCES

- [1] T.-Y. Oh *et al.*, “A 7 Gb/s/pin GDDR5 SDRAM with 2.5 ns bank-to-bank active time and no bank-group restriction,” in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 434–435.
- [2] T.-Y. Oh *et al.*, “A 3.2 Gb/s/pin 8 Gb 1.0 V LPDDR4 SDRAM with integrated ECC engine for sub-1 V DRAM core operation,” in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2014, pp. 430–431.
- [3] D. U. Lee *et al.*, “A 1.2 V 8 Gb 8-channel 128 GB/s high-bandwidth memory (HBM) stacked DRAM with effective microbump I/O test methods using 29 nm process and TSV,” in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2014, pp. 432–433.
- [4] Y. J. Yoon *et al.*, “An 1.1 V 68.2 GB/s 8 Gb wide-IO2 DRAM with non-contact microbump I/O test scheme,” in *IEEE ISSCC Dig. Tech. Papers*, Jan./Feb. 2016, pp. 320–322.
- [5] *High Bandwidth Memory (HBM) DRAM*, document JESD235A, JEDEC Solid State Technology Association, Nov. 2015.
- [6] D. U. Lee *et al.*, “An exact measurement and repair circuit of TSV connections for 128 GB/s high-bandwidth memory (HBM) stacked DRAM,” in *Symp. VLSI Technol. Dig. Tech. Papers*, Jun. 2014, pp. 1–2.

- [7] A. L. S. Loke *et al.*, “Loopback architecture for wafer-level at-speed testing of embedded HyperTransport processor links,” in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Sep. 2009, pp. 605–608.



Kyomin Sohn received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, South Korea, in 1994 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from KAIST, Daejeon, South Korea, in 2007.

From 1996 to 2003, he was with Samsung Electronics, South Korea, where he was involved in SRAM Design Team. He designed various kinds of high-speed SRAM devices, which are used for external cache and buffer memory. He rejoined Samsung Electronics, Hwaseong, South Korea, in 2007, where he has been involved in DRAM Design Team. His current research interests include stacked DRAM, such as HBM, robust memory design, and next-generation memory devices.

He has been serving as a Technical Program Committee Member of Symposium on VLSI Circuits since 2012.



Won-Joo Yun was born in Daejeon, South Korea, in 1978. He received the B.S. degree from Korea University, Seoul, South Korea, in 2000, the M.S. degree from Hanyang University, Seoul, in 2002, and the Ph.D. degree from Korea University in 2009, all in electrical engineering.

In 2004, he joined Hynix Semiconductor Inc., Ichon, South Korea, where he was involved in circuit design for high-speed DRAMs, such as GDDR3 SDRAM. From 2009 to 2012, he was with Keio University, Yokohama, Japan, as a Post-Doctoral Research Associate, where he was involved in inductive coupling and transmission line coupling interface. In 2012, he joined Samsung Electronics, Hwaseong, South Korea, where he has been involved in circuit design for DDR2 and DDR4, where he has been in charge of IO circuit design and DFX architecture for HBM. He has authored over 20 papers in the field of electronic circuits. He holds more than 40 U.S. patents. His current research interests include low-power high-speed interface circuit design, 3-D ICs, and design for excellence.



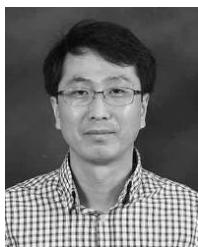
Reum Oh received the B.S., M.S., and Ph.D. degrees engineering from Korea University, Seoul, South Korea, in 1999, 2001, and 2015, respectively, all in electrical and electronics.

She joined Samsung Electronics, Hwaseong, South Korea, in 2001, where she was involved in circuit design for DDR2, DDR3, DDR4, and HBM. Her current research interests include TSV and high-bandwidth designs for 3-D and 2.5-D memory.



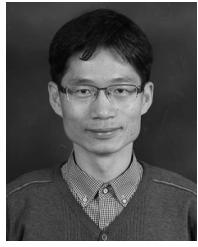
Chi-Sung Oh received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, South Korea, in 1966 and 1988, respectively.

In 1988, he joined Samsung Electronics, Hwaseong, South Korea, where he has been involved in mobile DRAM and Wide IO DRAM. His current research interests include the design of HBM.



Seong-Young Seo was born in Gunsan, South Korea, in 1973. He received the B.S. degree from Soongsil University, Seoul, South Korea, in 1999, and the M.S. degree from Hanyang University, Seoul, in 2001, and the Ph.D. degree from Sungkyunkwan University, Suwon, South Korea, in 2012, all in electrical engineering.

In 2001, he joined Samsung Electronics, Hwaseong, South Korea, where he is involved in the design of Rambus, DDR, and HBM DRAM for memory applications. He is currently a Principal Engineer in Samsung, where he is responsible for the design of DRAM data path and the at-speed test of HBM DRAM. His current research interests include the design of high-speed memory and high frequency measurement.



Min-Sang Park received the B.S. degree in electronic engineering from Ajou University, Suwon, South Korea, in 1996, and the M.S. degree in electric and electronic engineering from Pohang University of Science and Technology, Pohang, South Korea, in 1998.

He joined Samsung Electronics, Hwaseong, South Korea, in 1998, where he is currently with the DRAM Design Team. He has been involved in the design of GDDR, DDR, and HBM. His current research interests include robust memory design and

device thermal solutions.



Dong-Hak Shin was born in Iksan, South Korea, in 1970. He received the B.S. degree from Ajou University, Suwon, South Korea, in 1996, and the M.S. degree from Sungkyunkwan University, Suwon, South Korea, in 2002, all in electrical engineering.

In 1996, he joined Samsung Electronics, Hwaseong, South Korea, where he is involved in the design of SDR, Rambus, DDRx, and HBM, for high-speed and high-density core circuits. He has authored over ten U.S. patents. His current research interests include high-speed and low-power

bitline sense amplifier and core circuit design for high-density memory for excellence.



Won-Chang Jung was born in Daegu, South Korea, in 1973. He received the B.S. degree from Kyung-buk National University, Daegu, South Korea, in 1996, and the M.S. degree from Pohang University of Science and Technology, Pohang, South Korea, in 1998, all in electrical engineering.

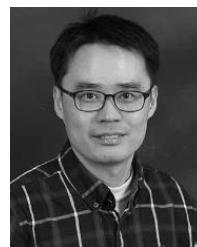
In 1998, he joined Samsung Electronics, Hwaseong, South Korea, where he is currently involved in the design of DDR2, DDR3 DRAM. He has been in charge of CORE design for HBM.

His current research interests include high-bandwidth memory, high-density memory, and 3D ICs.



Sang-Hoon Shin was born in Seoul, South Korea, in 1976. He received the B.S. and M.S. degree from Sogang University, Seoul, South Korea, all in electrical engineering.

In 2004, he joined Hynix Semiconductor Inc., Ichon, South Korea, where he was involved in circuit design for GDDR3 SDRAMs and 3DS DRAMs. In 2013, he joined Samsung Electronics, Hwaseong, South Korea, where he has been involved in circuit design. He has been in charge of DFX architecture and circuit design for HBM. He has authored more than five papers in the field of electronic circuits and he holds more than 15 U.S. patents. His current research interests include high-bandwidth memory, high-density memory, 3D ICs, and DFX.



Je-Min Ryu was born in Daejeon, South Korea, in 1976. He received the B.S. degree in electrical engineering from Hanyang University, Seoul in 2003.

In 2003, he joined Samsung Semiconductor Inc., Hwaseong, South Korea, where he was involved in circuit design for high-speed DRAMs, such as DDR3/DDR4/LP4. He has been in charge of repair circuit design and soft/hard Cell PPR/hard LANE architecture for HBM.



design for memory.

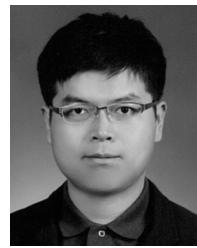
Hye-Seung Yu received the B.S. degree in semiconductor engineering from Dongguk University, Seoul, South Korea, in 2003, and the M.S. degree in electrical and computer engineering from Korea University, Seoul, South Korea, in 2005.

She joined Samsung Electronics, Hwaseong, South Korea, in 2005, where she has been involved in the I/O interface circuit design for DDR2, DDR3, DDR4, and HBM. Her current research interests include low-power and high-speed interface circuit



Jae-Hun Jung received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Chung-Ang University, Seoul, South Korea, in 2006, 2008 and 2013, respectively.

He joined Samsung Electronics, Hwaseong, South Korea, in 2013, where has been involved in the I/O interface circuit design for DDR3, DDR4, LPDDR3, LPDDR4, and HBM. His current research interests include high-speed frequency synthesizers, clock/data recovery circuits, and low-power robust interface circuit design for memory.



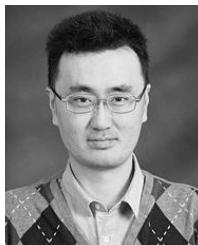
Hyunui Lee received the B.E. degree in computer engineering (major) and electronic engineering (minor) from Seokyeong University, Seoul, South Korea, in 2008, and the M.S. and Ph.D. degrees in physical electronics from the Tokyo Institute of Technology, Tokyo, Japan, in 2010 and 2013, respectively.

In 2013, he joined Samsung Electronics, Hwaseong, South Korea, where has been working on the I/O interface circuit design for DDR3, DDR4, and HBM. His current research interests include high-speed and high-resolution data converters, low-power robust interface circuit design for memory, and test circuit design, such as DFX.



Seok-Yong Kang received the B.S. degree in semiconductor system engineering from Sungkyunkwan University, Suwon, South Korea, in 2010.

He joined Samsung Electronics, Hwaseong, South Korea, in 2010, where has been involved in the I/O interface circuit design for DDR2, DDR3, DDR4, and DFX, high-speed serial-link test circuit for HBM. His current research interests include analog and digital circuit design for high-speed interface, design for excellence, and high-speed serial-link test.



Young-Soo Sohn received the B.S. degree from Sogang University, Seoul, South Korea, in 1997 and the M.S. and Ph.D. degrees from the Pohang University of Science and Technology, Kyungbuk, South Korea, in 1999 and 2003, respectively, all in electronic engineering.

He joined Samsung Electronics, Hwaseong, South Korea, in 2003, where he has been involved in developing high bandwidth DRAM, such as XDR, GDDR5, and HBM. His current research interests include high-speed CMOS circuit design, power/signal integrity, and interconnect modeling.



Seong-Jin Jang received the B.S. degree in electronic engineering from Kyungpook National University, Daegu, South Korea, in 1987, and the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Seoul, South Korea, in 1990.

He joined LG Semicon Corporation Ltd., Seoul, South Korea, in 1990, where he was involved in the DRAM Design Division. Since 2000, he has been with Samsung Electronics, Hwaseong, South Korea, as a Vice President of the DRAM Design Division. His current research interests include high-speed DRAM and interface design.



Jung-Hwan Choi was born in Daegu, South Korea, in 1968. He received the B.S. degree from Kyungpook National University, Daegu, South Korea, in 1990, and the M.S. and Ph.D. degrees from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1992 and 1997, respectively, all in electrical engineering.

In 1997, he joined Samsung Electronics, Hwaseong, South Korea, where he is involved in the design of Rambus, XDR DRAM, and high-speed I/O interface for memory applications. He is a Master in Samsung, where he is responsible for the design of DRAM interface and the development of high-speed DRAM interfaces for the next generation, including LPDDRx and DDRx. His current research interests include the design of monolithic microwave IC, high-speed memory, and high-frequency measurement.



Yong-Cheol Bae received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, South Korea, in 1992 and 1994, respectively.

In 1994, he joined Samsung Electronics, Hwaseong, South Korea, as a Research Engineer, where he has been involved in the circuit design of high-speed low-power DRAMs. In 2008, he joined the Mobile DRAM I/O Group and had been in charge of developing low-power interfaces of LPDDR2/3/4 and wide-I/O. Since late 2016, he has been leading the HBM Design Group to develop HBM Gen2 and beyond as a Vice President. His current research interests include the design of high-speed low-power memory and VLSI systems and I/O design for signal integrity of DRAM-AP channels.



Gyoyoung Jin was born in Seoul, South Korea, in 1962. He received the Bachelor's degree in electrical engineering and the Ph.D. degree from Seoul National University, Seoul, South Korea.

He is in charge of the DRAM Product and Technology of Memory Division, Samsung Electronics Company, where he is currently leading the state-of-the-art DRAM design and process technology. He has been a Samsung Fellow since 2011. He led the Memory Technology Development Team in Semiconductor Research and Development Center. His special focus area was in DRAM and Flash memory, where he led future generation 1 x nm process device development.