Multi-Slot Main Memory System for Post DDR3

Jaejun Lee, Sungho Lee, and Sangwook Nam, *Member, IEEE*

Abstract—This brief introduces a suitable architecture for a high-data rate and high-density system using bidirectional single-ended signaling. For chip-to-chip interconnections requiring high speed and high density for the main memory system, an SSTL-II-based structure was previously used. However, this structure is no longer applicable for higher speeds at higher densities. By using an optimum reflection coefficient at the junction of a branch, a multislot system acts in the same way as a point-to-point system. This architecture significantly improves the signal integrity. The simulated jitter and eye openings, including transmission line loss, were improved by 53.4% for write operation and 65.1% for read operations at 3.2 Gbps under heavy loading conditions. The peak-to-peak time jitters of 67.1 and 72.0 ps were measured at 3.3 Gbps.

Index Terms—Electric impedance, equalizers, interconnections, intersymbol interference (ISI), jitter, noise, routing.

I. Introduction

SINCE the performance of personal computers (PCs) has improved at an accelerated rate over the past decade, the demand for memory with a larger bandwidth and higher density has dramatically increased. At present, reliable communication between the memory controller and the main memory is one of the major challenges in this field due to the high operation frequency and the loading effect from the number of dual inline memory modules (DIMMs) attached to a channel. Therefore, multidrop buses such as the SSTL-II bus topology shown in Fig. 1 have traditionally been used as the main memory buses in PCs. However, as the data rates on these buses have increased, the maximum number of slots per channel declined in order to keep the signal integrity. The number of slots per channel has been reduced due to intersymbol interference (ISI) caused by signal reflections at impedance mismatches in multidrop junctions and terminations. Although the capacity per module has increased, the reduction of slots per channel limits the memory capacity per channel [1]. For these reasons, the current

J. Lee and S. Nam are with the Applied Electromagnetics Laboratory, School of Electrical Engineering and Computer Science, Institute of New Media and Communications, Seoul National University, Seoul 151-742, Korea (e-mail: jaejun@ael.snu.ac.kr).

S. Lee is with the Applied Electromagnetics Laboratory, School of Electrical Engineering and Computer Science, Institute of New Media and Communications, Seoul National University, Seoul 151-742, Korea, and also with the Korea Electronics Technology Institute, Seongnam 463-816, Korea.
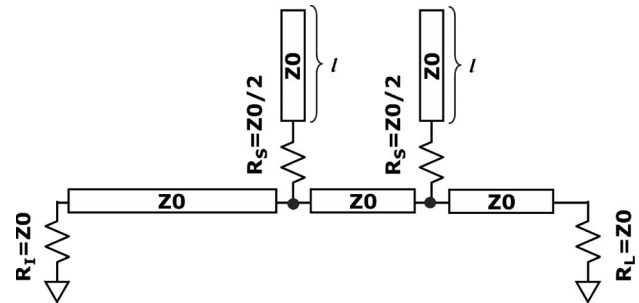
Fig. 1. SSTL-II bus topology.

DDR3 memory buses have only two slots per channel. In the near future, a type of point-to-point bus will prevail without any equalizer circuits in dynamic random-access memory (DRAM). Techniques such as parallel channels and fully buffered DIMM [1] are used to fulfill memory capacity demands. However, these are expensive in terms of input/output pins, PC board area, extra error correction circuits, and thermal issues. In addition, many papers have studied equalizer circuits used to eliminate distortion at the receiver block, such as the feedforward equalizer and the decision feedback equalizer [2], and these receiver equalizer schemes could be good solutions for solving the ISI problem; however, they are not suitable for low-power cost-effective systems, such as PCs.

This brief introduces a new type of multidrop bus for the main memory channel that acts as a point-to-point bus. From the basic two-slot solution, an extension method for a three-slot bus is suggested with an on-die-termination (ODT) on/off configuration. The proposed multidrop bus scheme improves the signal integrity at higher operating frequency, thereby reducing the total jitter and widening the eye openings. In addition, this system reduces the specification burden of the driver and the receiver for post DDR3 for which the target operating data rate should cover from 1.6 to 3.2 Gbps without any equalizer. This method can also be applied to high-speed signaling through a parallel link.

II. Design of Architecture

The basic concept of the new multidrop bus topology is to minimize the reflection noise from the channel like a point-to-point bus topology. The conventional multidrop bus topologies that have been used in DRAM channels are based on the SSTL-II shown in Fig. 1. These are good enough for relatively low frequencies. However, the current DDR3 bus topology based on the SSTL-II has structural weak points for practical applications in high-frequency operations. Fig. 2 shows the current DDR3 main memory system architecture. The gap A
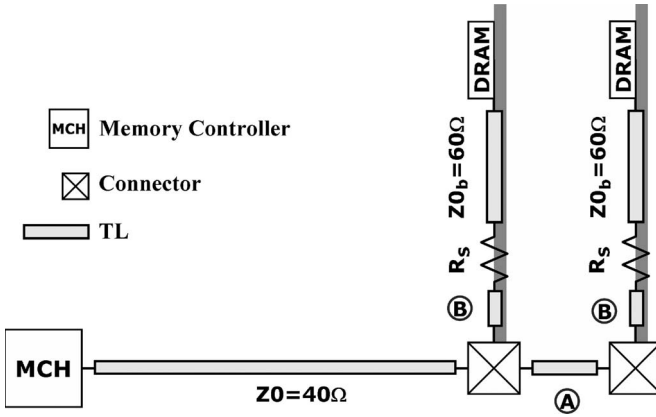
Fig. 2. DDR2 and DDR3 main memory bus system with two slots.

TABLE I
ODT CONFIGURATION FOR THE DDR3 MEMORY SYSTEM
(DOUBLE-RANK TWO-SLOT CASE [4])

Operation Status	ODT Values (Ω)		
	MCH	1st Slot	2nd Slot
Write to 1st Slot	X	60	20
Write to 2nd Slot	X	20	60
Read from 1st Slot	50	X	20
Read from 2nd Slot	50	20	X

* Z0 of PC board is 40Ω, and Z0 of DIMM is 60Ω

of the slots and the length B of the junction to a resistor prevent a perfect matching at the multidrop junction, as shown in Fig. 2. For these reasons, residual reflection noise exists in the channel. Therefore, for operations over 1.6 Gbps, the current DDR3 bus structure is difficult to use for a post-DDR3 memory system channel without any equalizer circuits.

In the current DDR3 main memory system, the maximum number of slots is limited to two. When this topology is used over 1.6 Gbps, the only solution is to use a one-slot memory system or implement the various equalizing schemes in DRAMs. If an equalizer is applied to a DRAM, issues about power consumption and die area arise. Moreover, the present DDR3 main memory architecture consumes routing space on the PC board, as the 40-Ω impedance traces used on PC boards to improve the signal integrity for the write operation are about two times wider than the traces with an impedance of 60 Ω. All DRAMs must have an ODT with a complicated configuration and different kinds of values for ODT options, as shown in Table I [4]. This is one of the design issues, and the current flowing in these ODTs might also cause thermal issues. In the post DDR3 memory system that works at over 1.6 Gbps, a simpler channel structure without lumped elements at the multidrop junctions is needed.

A. Concept of the New Multidrop Bus Topology

As Fig. 3 shows, when the impedance of the branch transmission line is half of the main impedance Z0, the input impedance
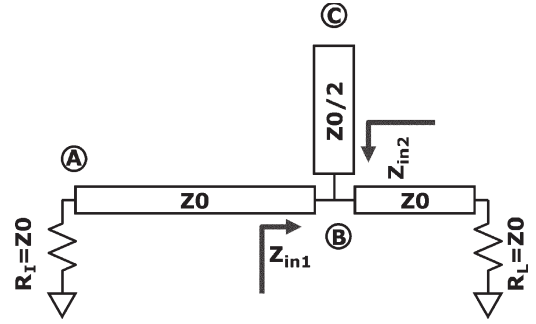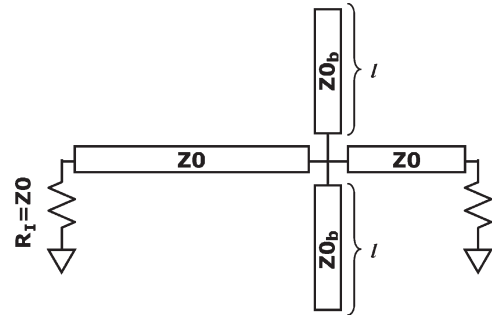

Fig. 3. Concept of the proposed topology.

Fig. 4. Example of the equivalent topology to the concept of the proposed topology.

Z_{in1} is Z0/3 by

$$Z_{in1} = \frac{Z0}{2} // Z0 = \frac{Z0}{3} \quad \Gamma_{b1} = \frac{Z0/3 - Z0}{Z0/3 + Z0} = -\frac{1}{2} \quad (1)$$

meaning that the reflection coefficient Γ_{b1} is equal to $-1/2$. Thus, the voltage at Point B is half of the incident voltage at Point A. Since Point C is open, the open-circuit voltage at this point is twice the input voltage. This means that the voltage of Point C is same as the incident voltage of Point A.

The reflected signal at Point C of Fig. 3 does not come back to Point C again, because Z_{in2} is Z0/2 and Γ_{b2} is 0 by

$$Z_{in2} = \frac{Z0}{2} \quad \Gamma_{b2} = \frac{Z0/2 - Z0/2}{Z0/2 + Z0/2} = 0. \quad (2)$$

From the conceptual bus topology shown in Fig. 3, the relationship

$$Z0_b = \frac{Z0}{2} \cdot N \quad (3)$$

between the impedance of the branch $Z0_b$ and the number of the branches N can be derived when all of the lengths of the branches are the same.

For example, in the case of a two-drop bus topology $(N = 2)$, both the main transmission line and the branch trace have the same impedance $(Z0 = Z0_b)$, by (3), as shown in Fig. 4. The topology in Fig. 4 is exactly equivalent to the proposed conceptual bus topology in Fig. 3 in its electrical characteristics. From (3), the proposed structure seems to be applicable for many branches; however, two or three branches (stubs) may be the maximum number, and the topology with two branches may be the best choice, taking productivity into consideration.
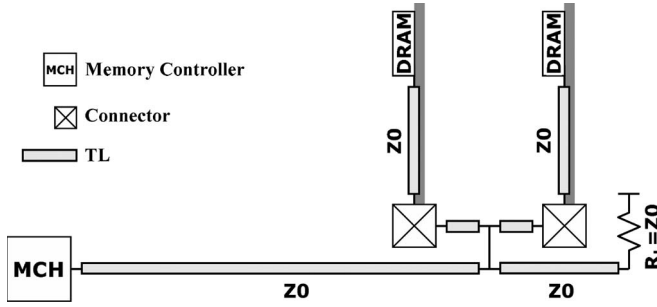
Fig. 5.   Proposed main memory bus system with two slots.

TABLE II
ODT CONFIGURATION FOR THE PROPOSED TWO-SLOT
MAIN MEMORY ARCHITECTURE

Operation Status	ODT Values (Ω)		
	MCH	1st Slot	2nd Slot
Write to 1st or 2nd Slot	X	X	X
Read from 1st Slot	75	X	50
Read from 2nd Slot	75	50	X

* Z0 of the PC board and DIMM is 50Ω

B. Application to Main Memory Bus Topology

Fig. 5 shows the new architecture of the main memory system with the proposed topology in Fig. 4. As shown in Fig. 4, the trace of DIMM impedance Z_{0b} is the same as the impedance Z0 of the PC board. This characteristic makes it easy to produce and apply the multidrop bus system, which saves the routing area for the PC board, because 50- or 60-Ω impedance can be chosen, instead of 40-Ω trace impedance for the DDR3 topology. As shown in Table II, the ODT control scheme and values are simplified. In the write operation case, the proposed memory system does not require ODT. Since the current does not flow to the operating DRAM, thermal issues in the DRAM are mitigated. In the proposed memory system scheme, the trace length of DIMM does not affect the signal integrity, but it does contribute to attenuation only as long as the lengths of the branches are the same. However, the proposed structures increase the routing density in the slot-to-slot area.

To perfectly construct the proposed channel, the end of the slot acts as an open end for all frequencies. However, the DRAM receiver has an input capacitance of about 1.5 pF, including the driver output capacitance, electrostatic-discharge capacitance, package, etc. This unwanted parasitic capacitance shows a small reactance during high-frequency operations. This effect is critical in heavy loading condition; in other words, all populated DIMMs have double ranks in one channel, as shown in Fig. 6. To resolve the degradation of signal integrity during heavy loading, a series inductor L_S (2.5 nH) is attached in front of the DRAM. Since this series inductor L_S and parallel input capacitance work as a lumped transmission line, the end of the stub acts as an open end. In the case of the single-rank DIMM shown in Fig. 5, a series inductor in front of the DRAM is not required, because the inductance of the DRAM package works as a series inductor.
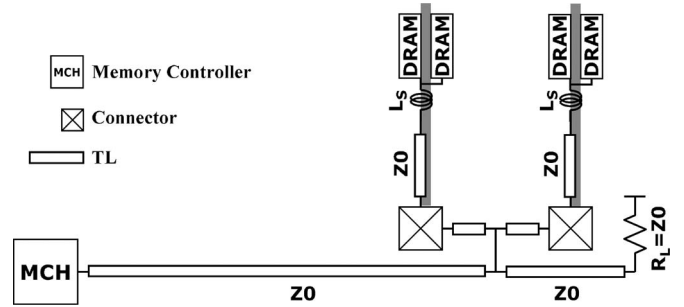

Fig. 6.   Proposed architecture with heavy loading condition.
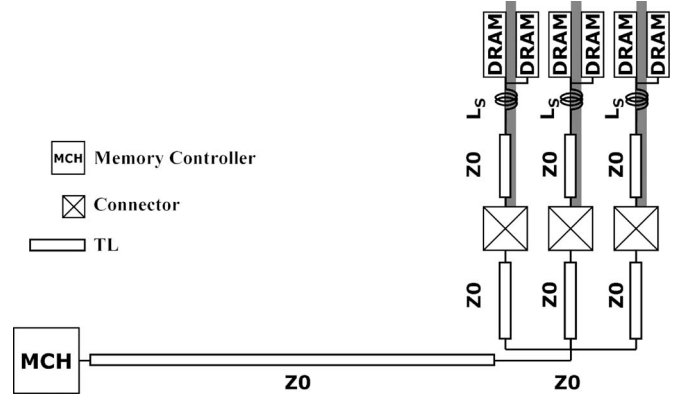

Fig. 7.   Proposed architecture of the three-slot memory system.

TABLE III
ODT CONFIGURATION FOR THE PROPOSED THREE-SLOT
MAIN MEMORY ARCHITECTURE

Operation Status	ODT Values (Ω)			
	MCH	1st Slot	2nd Slot	3rd Slot
Write to 1st or 2nd Slot	X	X	X	50
Write to 3rd Slot	X	X	50	X
Read from 1st Slot	75	X	50	50
Read from 2nd Slot	75	50	X	50
Read from 3rd Slot	75	50	50	X

* Z0 of PC board and DIMM is 50Ω

In the case of the three-slot memory system, there are two possible ways to construct the memory system. One is to use the three-branch topology of Z_{0b} with 60 Ω and Z0 with 40 Ω by (3). This three-branch topology has a routing issue that comes from the wide-trace-width 40-Ω transmission line as a conventional DDR3 topology. However, when using on/off switchable ODT, a three-slot system can consist of the proposed two-branch topology, as displayed in Fig. 3. The architecture for a three-slot memory system based on two-branch topology is shown in Fig. 7. The ODT values and configuration are controlled by data in Table III. In this way, we found the three-slot memory system solution.

In the proposed topology, the signal integrity is sensitive to the loading balance since the conceptual topology shown in Fig. 2 must be maintained. When only one slot is populated in the proposed two-slot system, the signal integrity is worse
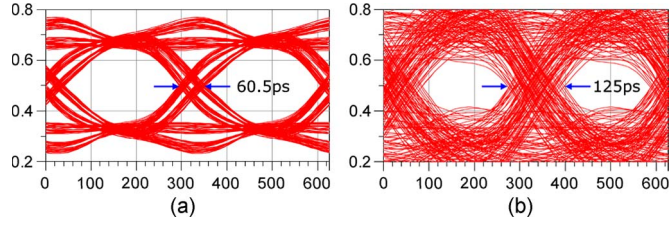
Fig. 8. (a) Eye diagram of write operation and (b) eye diagram of read operation for the DDR3 architecture in two-slot heavy loading case at 3.2 Gbps.
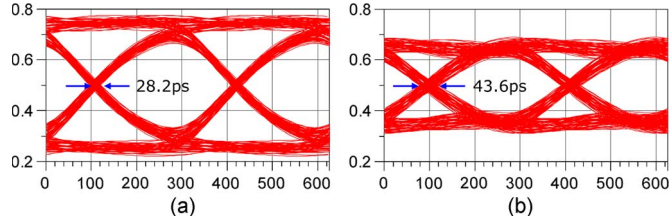

Fig. 9. (a) Eye diagram of write operation and (b) eye diagram of read operation for the proposed architecture in two-slot heavy loading case at 3.2 Gbps.

TABLE IV
SIMULATION RESULTS OF MAIN MEMORY ARCHITECTURES

Architecture Type	Write Operation			Read Operation		
	Jitter (P-P)	Jitter (RMS)	Eye Height	Jitter (P-P)	Jitter (RMS)	Eye Height
DDR3 (SSTL-II)	60.5ps	20.3ps	277mV	125ps	30.8ps	166mV
Proposed 2-slot	28.2ps	6.0ps	288mV	43.6ps	11.9ps	201mV
Proposed 3-slot	49.3ps	11.0ps	323mV	60.5ps	14.5ps	181mV

* All DIMMs are double ranks (heavy loading condition) in 3.2Gbps
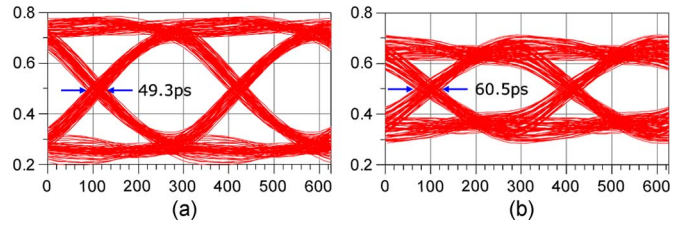

Fig. 10. (a) Eye diagram of write operation and (b) eye diagram of read operation for the proposed architecture in three-slot heavy loading case at 3.2 Gbps.

III. SIMULATIONS AND MEASUREMENTS

A. Simulated Results

To validate the proposed method, simulations were performed for the multislot memory system with an FR-4 substrate ε_r of 4.1, which is generally used in PC boards and DIMMs. The connector model was the current DDR3 through-hole DIMM connector model. The Advanced Design System of Agilent Technologies was used to simulate the signal integrity of the memory system. For the simulation, a V_{DD} of 1.0 V and ideal voltage sources with a source impedance of 50 Ω were used. The DRAM package model was the same as the DDR3 BGA package model. The length of the memory controller for the first connector was set to 4 in. The trace impedance of the proposed architecture was 50 Ω for both the PC board and DIMM, whereas the conventional PC board impedance was 40 and 60 Ω for the DIMM and for the present DDR3, respectively. Figs. 8 and 9 show the eye diagrams produced for a double-rank DIMM in all populated two-slot systems for the conventional DDR3 memory system structure and the proposed structure at 3.2 Gbps, respectively. As shown in Table IV, the rms jitter of the proposed structure was improved by 53.4% and 65.1% for writing and reading, respectively.

The proposed three-slot memory system showed signal integrity that is worse than that of the proposed two-slot solution. However, the proposed three-slot system still had an improved timing margin and voltage margin, compared with the conventional scheme, as shown in Fig. 10 and Table IV. Notice that it has higher density than the conventional DDR3 two-slot and the eye height is also improved. A significantly reduced jitter
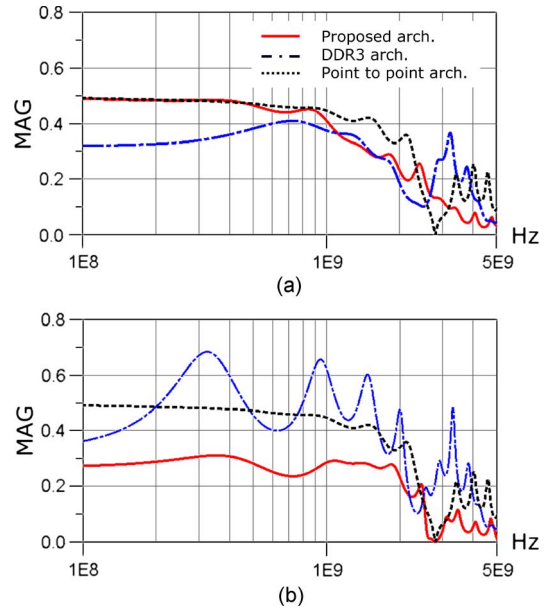

Fig. 11. (a) Frequency response of write operation and (b) frequency response of read operation for the architectures in heavy loading case.

improves the timing margin when the memory system reads from DRAM and writes to DRAM.

Fig. 11 shows the frequency response of the proposed architecture and reveals a relatively flat magnitude response from low frequency to 1.6 GHz, just like the point-to-point topology. This means that the preemphasis driver will not be required until 3.2 Gbps, even though it depends on the channel length. If the preemphasis driver that is needed to overcome the high-frequency attenuation is used, this proposed topology might be extended to over 3.2 Gbps.
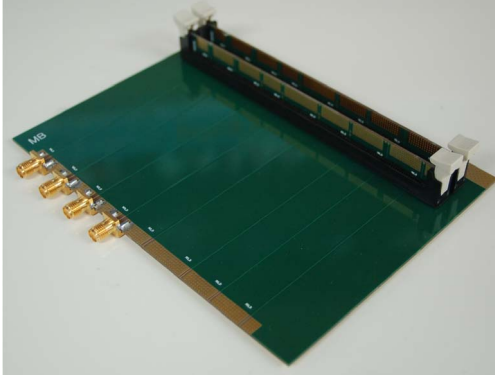
Fig. 12.   Test board.

B. Measured Results

The proposed architecture of the two-slot memory system was implemented to check the signal integrity at 3.2 Gbps. A simplified PC board and DIMMs having a 4-in 50-Ω trace for the PC board and 0.5-in 50-Ω trace from the tab of the DIMM to the capacitor load for the DIMM were fabricated with 4.1-ε_r four-layer PCB, as shown in Fig. 12. A conventional DDR2 through-hole connector was used to connect the PC board to the DIMM. Drivers for the memory controller and DRAMs were substituted for with the Tektronix AWG7102 signal generator. Because of the sampling limitation of AWG7102, the driving signal has the limited wave of 1.0 V_{p-p}, a 100-ps rise/fall time, and a 600-ps pulsewidth at 3.3 Gbps. This data rate was somewhat higher than the targeted data rate of 3.2 Gbps, but it was enough to show the performance of the architecture. We assumed that the receiver works as capacitor load that is electrically equivalent to the input capacitance. Therefore, 1.5-pF capacitor loads were mounted. The eye diagram measurements were made with a digital oscilloscope Tektronix DSA 72004 B and high-frequency probe P7513 A.

As shown in Fig. 13, the jitter of 67.1 ps and the eye height of 172 mV were measured in the write operation. In addition, 72.0-ps jitter and 148-mV eye height were measured in the read operation. Compared with the simulation in Fig. 9 and Table IV, the experiment showed to be much larger than the simulated jitter. The experimental measured jitters include the driver jitter and noise from the SMA connector, cable, etc. It is believed that the DDR2 connector has a lower impedance (about 30 Ω) than the 50-Ω trace impedance, so that the reflection noise from the connector mainly affected the jitter in the experiment. In this test fixture, the signal pins in connectors were surrounded by ground pins to eliminate unwanted noise, so that it made the impedance of the connectors decrease and deteriorated the eye diagram in Fig. 14. It shows more severe effects in the write operation than in the read operation since all the ends of interconnect were terminated by ODT in the read operation. To enhance the jitter characteristics for the post DDR3 system, impedance-controlled connectors were strictly required.

IV. CONCLUSION

In this brief, an effective architecture for the main memory system has been proposed to minimize the reflection noise and
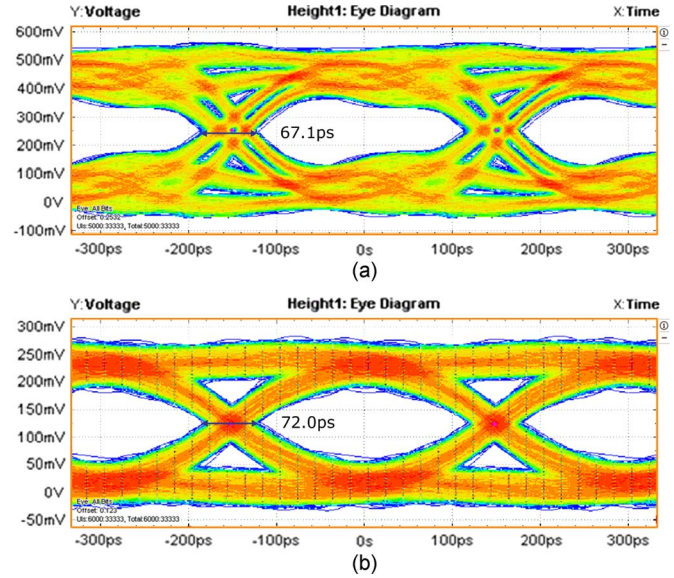

Fig. 13.   (a) Measured eye diagram of write operation and (b) measured eye diagram of read operation for the proposed architecture in two-slot heavy loading case at 3.3 Gbps.
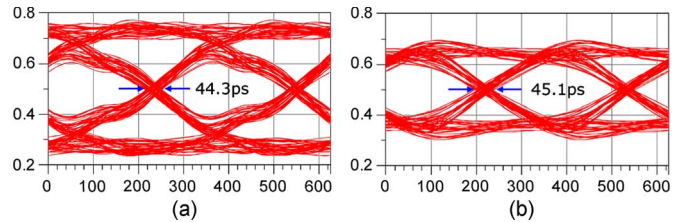

Fig. 14.   (a) Simulated eye diagram of write operation and (b) eye diagram of read operation with low impedance connector for the proposed architecture in two-slot heavy loading case at 3.3 Gbps.

extend its use to higher frequency with multislots. As a result, the system has obtained better jitter and eye height than the conventional main memory structure based on SSTL-II. Since the current does not flow to the operating DRAM, it is expected to diminish the thermal problems in the present DDR3 memory module. Therefore, the proposed architecture can be used as a high-frequency and high-density solution for post-DDR3 main memory systems.

REFERENCES

[1] J. Haas and P. Vogt, Fully-Buffered DIMM Technology Moves Enterprise Platforms to the Next Level, Mar. 2005.
[2] H. Fredriksson and C. Svensson, "2.6 Gb/s over a four-drop bus using an adaptive 12-tap DFE," in *Proc. Eur. Solid-State Circuits Conf.*, Sep. 2008, pp. 470–473.
[3] H. Chung, Y. Jang, Y. Choi, H. Park, J. Kim, S. Lim, J. Sunwoo, M. Park, H. Kim, S.-Y. Kim, H.-K. Kim, S.-J. Chung, E.-M. Lee, Y. Kim, Y.-S. Lee, W.-S. Kim, J.-B. Lee, and C. Kim, "Channel BER measurement for a 5.8 Gb/s/pin unidirectional differential I/O for DRAM application," in *Proc. IEEE Asian Solid-State Circuits Conf.*, Nov. 2008, pp. 29–32.
[4] *DDR3 SDRAM Standard (JESD 79-3B)*, JEDEC Solid State Technol. Assoc., Arlington, VA, Apr. 2008, pp. 91–108.
[5] M. S. Sharawi and M. T. Al-Qdah, "The design and simulation of a 400/ 533 Mbps DDR-II SDRAM memory interconnect bus," in *Proc. IEEE Int. Multi-Conf. Syst., Signals Devices*, Jul. 2008, pp. 1–6.
[6] W. T. Beyene and A. Amirkhany, "Controlled intersymbol interference design techniques of conventional interconnect systems for data rates beyond 20 Gbps," *IEEE Trans. Adv. Packag.*, vol. 31, no. 4, pp. 731–740, Nov. 2008.