A 1.2-V 1.5-Gb/s 72-Mb DDR3 SRAM

Cho, Uk Rae; Kim, Tony Tae-Hyoung; Yoon, Yong-Jin; Lee, Jong Cheol; Bae, Dae Gi; Kim, Nam Seog; Kim, Kang Young; Son, Young Jae; Yang, Jeong Suk; Sohn, Kwon Il; Kim, Sung Tae; Lee, In Yeol; Lee, Kwang Jin; Kang, Tae Gyoung; Kim, Su Chul; Ahn, Kee Sik; Byun, Hyun Geun

2003

Cho, U. R., Kim, T. H., Yoon, Y. J., Lee, J. C., Bae, D. G., Kim, N. S., et al. (2003). A 1.2-V 1.5-Gb/s 72-Mb DDR3 SRAM. IEEE Journal of Solid State Circuits, 38(11), 1943-1951.

<https://hdl.handle.net/10356/100928>

<https://doi.org/10.1109/JSSC.2003.818137>

© 2003 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder. <http://www.ieee.org/portal/site> This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

A 1.2-V 1.5-Gb/s 72-Mb DDR3 SRAM

Uk-Rae Cho, Tae-Hyoung Kim, Yong-Jin Yoon, Jong-Cheol Lee, Dae-Gi Bae, Nam-Seog Kim, Kang-Young Kim, Young-Jae Son, Jeong-Suk Yang, Kwon-Il Sohn, Sung-Tae Kim, In-Yeol Lee, Kwang-Jin Lee, Tae-Gyoung Kang, Su-Chul Kim, Kee-Sik Ahn, and Hyun-Geun Byun

Abstract—A 1.2-V 72-Mb double data rate 3 (DDR3) SRAM achieves a data rate of 1.5 Gb/s using dynamic self-resetting circuits [5]. Single-ended main data lines halve the data line precharging power dissipation and the number of data lines. Clocks phase shifted by 0° , 90° , and 270° are generated through the proposed clock adjustment circuits. The proposed clock adjustment circuits make input data sampled with optimized setup/hold window. On-chip input termination with the linearity error of $\pm 4.1\%$ is developed to improve signal integrity at higher data rates. A 1.2-V 1.5-Gb/s 72-Mb DDR3 SRAM is fabricated in a $0.10\text{-}\mu\text{m}$ CMOS process with five metals. The cell size and the chip size are $0.845\text{ }\mu\text{m}^2$ and 151.1 mm^2 , respectively.

Index Terms—1.5 Gb/s, 72 Mb, CMOS memory circuits, DDR3, high-speed SRAM, on-chip termination.

I. INTRODUCTION

HIGH-SPEED SRAMs are used in ultrafast systems, such as networking systems, servers and workstations. The development of these systems has dramatically increased the performance requirements of high-speed SRAMs. The main features asked on high-speed SRAMs are higher data rate and higher density, which are directly related to system performance, but it is quite difficult to implement SRAM with high density and high data rate. In high-speed SRAM, 32 Mb has been considered as the limit due to the chip size, standby current, speed, etc.

This paper describes a 1.2-V 1.5-Gb/s 72-Mb double data rate 3 (DDR3) SRAM which satisfies the data rate and the density required in recent ultrafast systems. The 72-Mb SRAM is fabricated using a $0.1\text{-}\mu\text{m}$ CMOS technology with the implemented SRAM, breaking the $1\text{-}\mu\text{m}^2$ barrier for the cell size. Some papers have been published including six-transistor (6T) SRAM cells near the $1\text{-}\mu\text{m}^2$ barrier [1]–[3]. One paper described a 6T embedded SRAM cell breaking the $1\text{-}\mu\text{m}^2$ barrier with the cell size of $0.998\text{ }\mu\text{m}^2$ [4]. Fig. 1 shows the trend in SRAM size recently reported. The implemented cell size of this work is $0.845\text{ }\mu\text{m}^2$, which is the smallest size to date. The chip size is 151.1 mm^2 . To halve the data line precharging power dissipation and the number of data lines, single-ended main data lines (SMDLs) are designed using dynamic self-resetting circuits [5]. The SRAM operates in two user-selectable modes: clock-aligned mode (CA mode) and clock-centered mode (CC mode). In CA mode, clocks phase shifted by 0° , 90° , and 270° from the external clock are generated through the proposed clock adjustment circuits to

Fig. 1. Trend of SRAM cell size reported.

sample address, control, and input data. In CC mode, clocks synchronized with the rising and falling edge of the external clock are generated and used to sample input signals. On-chip input termination with the linearity error of $\pm 4.1\%$ is developed to improve signal integrity at higher data rate. The impedance of input termination is programmable by external resistor R_T . A programmable impedance controller (PIC) tracks the process, voltage, and temperature (PVT) variations of the termination impedance.

The detailed architecture is described in Section II. In Section III, the SMDL scheme is presented. The proposed clock adjustment circuit is explained in Section IV. On-chip input termination is explained in Section V. Finally, the conclusions and hardware results are presented in Section VI.

II. CHIP ARCHITECTURE

Fig. 2 briefly shows a simplified architecture of the 72-Mb DDR3 SRAM. The SRAM is configured as either $2\text{M} \times 36$ or $4\text{M} \times 18$. The SRAM is divided into four mats each having nine I/Os, and each mat is subdivided by four submats for double-data-rate and burst-mode operations. These submats consist of 32 blocks, 16 blocks in the left and 16 blocks in the right, sharing section wordlines and wordline drivers. Each block has nine I/Os and each I/O is divided into 512 wordlines by 32 columns. Two submats are activated in each mat, and one block is activated in each selected submats. That is, eight blocks are accessed in four mats at the same time to provide the 72 bits required in double-data-rate operation.

The wordline decoder is composed of a wordline predecoder, main wordline decoders, and section wordline decoders. The wordline predecoder is located in the center of the chip and the main wordline decoders are located in the vertical channels between mat A and mat B and between mat C and mat D. Finally, the section wordline decoders are located in the center of each

Manuscript received April 9, 2003; revised June 23, 2003.

The authors are with the SRAM Memory Division, Samsung Electronics, Gyeonggi-Do 445-701, Korea (e-mail: purekth.kim@samsung.com).

Digital Object Identifier 10.1109/JSSC.2003.818137

Fig. 2. Architecture of 72-Mb DDR3 SRAM.

block. The bitline decoder consists of a bitline predecoder and main bitline decoders. The bitline predecoder is in the center of the chip and the main bitline decoders are distributed at the bottom of each block. The main data lines are distributed in four mats and divided into two stages. The first main data lines are located in the horizontal channels between submat X and submat Z, and submat Y and submat W. The second main data lines are routed vertically across the center of each mat to make balanced timing and minimum capacitance [6]. All I/O circuits and control circuits are located in the horizontal channel in the center of the chip.

III. SMDL

A. SMDL Scheme

In high-speed SRAMs such as the 4M DDR SRAM, differential main data lines have been used to reduce the delay through data line [7]. In the differential main data line scheme, one of the two differential data lines is precharged regardless of the data after finishing operation. Furthermore, the main data lines have large capacitive loads, causing large precharging power dissipation. In the SMDL scheme, the data line precharging operation is executed when the data is only "0". If the probability of "0" is equal to that of "1", SMDL reduces data line precharging power dissipation and the number of main data lines by half. Fig. 3 shows the SMDL scheme for an I/O in a mat. The outputs of 16 sense amplifiers in the near or far area are connected to one of four first main data lines (MDL1) and four MDL1s are connected to second main data lines (MDL2) through nMOS switches and drivers. The sense amplifier outputs of two submats are tied together because they are not selected at the same time in any burst-operation situations. Only one block is selected in submats W and Y. The other block is selected in submats X and Z.

When one out of 64 blocks is activated in two submats, the output of the sense amplifier in the selected block is transferred to MDL2 through nMOS switches and drivers. After that, the sampling clock RS_KCORE samples the data of MDL2 and transmits it to the data output buffer DLATCH. MDL1s, MDL_SUM, and MDL2 are initially precharged to V_{dd} . If the selected sense amplifier output is "0", one MDL1 connected

to the selected sense amplifier falls to Gnd while the other MDL1s remains at V_{dd} . If one of four MDL1s falls to Gnd, MDL_SUM also becomes Gnd. After T_{RS} from the falling edge of MDL_SUM, the reset signal turns the pMOS on and MDL_SUM is precharged to V_{dd} by a dynamic self-resetting circuit [5]. MDL_SUM can be precharged with the pulsewidth of T_{RS} because it has a small capacitive load compared to the main data lines. On the contrary, if the selected sense amplifier output is "1", the MDL1 connected to the selected sense amplifier as well as the other MDL1s remain at V_{dd} . In this case, four MDL1s turn the nMOS drivers off, causing MDL_SUM and MDL2 to remain at V_{dd} . That is, all nodes are in their initial states without additional power dissipation for precharging of MDL1s, MDL_SUM, and MDL2 when transmitting "1" to DLATCH. Precharging MDL1s, MDL_SUM, and MDL2 occurs only when the sense amplifier output data is "0". If the probability of "0" is equal to that of "1", power dissipation is reduced by half. In addition, the number of data lines is also reduced by half. Power dissipation in main data lines is reduced from 173 to 80 mW by adopting the SMDL scheme.

B. Sampling Clock (RS_KCORE) Generator

Another important point in SMDL is the timing of the sampling clock RS_KCORE. Since the data window of MDL2 is just T_{RS} , controlling the timing of RS_KCORE is very important to transfer the right data to DLATCH. In the conventional design, the sampling clock for DLATCH is derived from the external clock by adding delays. Since the data path is independent of the clock path, their timing difference becomes sensitive to PVT variations. The amount of timing difference caused by PVT variations can be neglected in low data rates. In high data rates, however, an automatic internal sampling clock generator, which is robust over PVT variations, is required.

As shown in Fig. 4, the sampling clock generator is composed of the same circuits used in SMDL except that it is differential. One I/O out of nine I/Os has differential outputs. The differential outputs of one sense amplifier in a selected block are used to generate the sampling clock. The operation is similar to that of SMDL. In precharge state, differential outputs of sense amplifiers are both precharged to V_{dd} , but in the access state, the differential outputs of the sense amplifier in a selected block become complementary. The sampling clock is generated from the moment when the outputs of the selected sense amplifier start to be complementary. The complementary sense amplifier outputs make MDL_SUM_T and MDL_SUM_C complementary. This makes the MDL2_T_C and RS_KCORE signal go high. After T_{KCORE} from the moment when RS_KCORE becomes complementary, the path from MDL2_T_C to RS_KCORE is disconnected and RS_KCORE becomes low. EN is an enabling signal indicating read operation. In this way, a sampling clock having pulsewidth of T_{KCORE} is generated. Like SMDL, one of MDL_SUM_T or MDL_SUM_C is precharged to V_{dd} after T_{RS} from the moment when MDL_SUM_T and MDL_SUM_C become complementary. The timing between RS_KCORE and reset signals is important because they are pulsed signals. For RS_KCORE to sample the data safely, data in MDL2 needs a setup margin. The timing difference between MDL2 and

Fig. 3. Schematic of SMDL scheme.

Fig. 4. Schematic of sampling clock (RS_KCORE) generator.

RS_KCORE plays the role of a setup time. To obey the setup time, data must arrive at MDL2 earlier than at RS_KCORE. In addition, the pulsewidth of RS_KCORE must be guaranteed to have the time for sampling. If the reset signal (/reset_T, /reset_C) is enabled while RS_KCORE is still sampling MDL2, the pulsewidth of RS_KCORE is reduced. The reduced pulsewidth of RS_KCORE can cause sampling to fail due to insufficient sampling time. To prevent this, T_{RS} must be larger than T_{KCORE} to guarantee the pulsewidth of RS_KCORE, T_{KCORE} . The delay from the sense amplifier output to MDL_SUM_T and MDL_SUM_C is the same as that from MDL1 to MDL_SUM in SMDL because the same circuits and the same layout architectures are used. Only the control circuit delay from four-input NAND gate to RS_KCORE is affected by PVT variations. The timing variances of the sampling clock are

minimized because the delay from the four-input NAND gate to RS_KCORE is relatively small compared with that of the sense amplifier output to MDL_SUM_T and MDL_SUM_C. In this sampling clock generator, the timing of the sampling clock is made highly correlated with that of the data. As a result, the timing of RS_KCORE becomes robust over PVT variations. Fig. 5 shows the timing of SMDL.

IV. CLOCK ADJUSTMENT CIRCUIT (CAC)

A. Timing Diagram of DDR3

The DDR3 SRAM supports two user-selectable modes: clock-centered (CC) mode and clock-aligned (CA) mode. In CC mode, all input signals are clock centered. Therefore, internal clocks synchronized with the rising and falling edges

Fig. 5. Timing of SMDL scheme.

Fig. 6. Input timing diagram of DDR3 SRAM and internal clocks in CA mode.

of the external clock are needed to sample input signals. In CA mode, address and control signals are clock centered, but input data is clock aligned. Therefore, clocks phase shifted by 90° and 270° are needed to sample input data with the optimized setup/hold window. Fig. 6 shows the timing diagram of the DDR3 (CA mode) input signals. Clocks for address, clock1, and clock2 are generated from clock adjustment circuits by shifting the phase of the input clock. The clock for address is synchronized with the rising edge of the external clock. Clock1 and clock2 are generated from the rising and falling edges of the external clock, shifted in phase by 90° . Two clock adjustment circuits are used to process the rising and the falling edges.

B. Clock Adjustment Circuit (CAC)

Fig. 7 shows the principle of the multiphase shift scheme. The input clock goes through two paths: the fast path and the delayed path, which consists of a delay block (Td) and a forward delay chain (FWD). A phase comparator compares the phases of the two paths and decides the length of FWD where the phase difference becomes 360° . If the delay Tm is added in both paths, the delay of FWD causing a phase difference of 360° becomes $T_{clk} - T_d + T_m$. The delayed clock goes through backward delay chain (BWD), which has the same delay as FWD. The

Fig. 7. Principle of multiphase shift.

Fig. 8. Block diagram of the proposed clock adjustment circuit.

total delay is $T_d + 2(T_{clk} - T_d + T_m) + T_d = 2T_{clk} + 2T_m$. This means that the output clock phase shifted by $2T_m/T_{clk}$ is generated after two clock cycles. In [8], T_m is zero. Therefore, only a clock synchronized with an external clock could be made, but, in the proposed CAC, T_m is not zero and the PVT variations of T_m are automatically compensated. By controlling the amount of the delay T_m , clocks phase shifted by any degrees can be generated.

Fig. 8 shows the block diagram of the proposed CAC. The CAC is composed of clock receivers (CLK_RCV), detection circuit mirror delay (DT_Mirror), JTAG controller (JTAG_CONT), clock driver (CLK_DRV), delay chain for 0° , and delay chain for 90° . The input clock goes through the shorter path to make the clock for address and control sampling. The total delay of the clock for address and control is

$$2T_{clk} = T_{d1} + T_{d1} + T_{d4} + T_{d3} + T_{d2} + 2\{T_{clk} - (T_{d1} + T_{d2} + T_{d3} + T_{d4})\} + T_{d4} + T_{d3} + T_{d2}.$$

That is, the internal clock synchronized with the external clock is generated in two clock cycles, as described in [8]. The input clock goes through the proposed longer path to make the clock for data. The total delay of the clock for data is

$$4T_{clk} + \frac{T_{clk}}{4} = T_{d1} + T_{d1} + T_{d4} + T_{d3} + T_{d2} + 2\{T_{clk} - (T_{d1} + T_{d2} + T_{d3} + T_{d4})\} + T_{d4} + T_{d3} + T_{d4} + 2(T_{clk} + T_{clk}/8 - T_{d4}) + T_{d4} + T_{d2}$$

which means that internal clock phase shifted by 90° is generated after four clock cycles. T_m in Fig. 6 is $1/8 T_{clk}$ in the proposed clock adjustment circuits. DT_Mirror is added in the delayed path to compensate for the delay of the phase comparators in the delay chains. CLK_RCV in the delayed path is to

Fig. 9. Block diagram of the delay chain for 90° phase shift.

compensate for the delay of CLK_RCV receiving the external clock.

Fig. 9 shows how the proposed delay chain for 90° is constructed. As mentioned before, a clock phase shifted by 90° is generated by adding delay units D in both the fast path and the delayed path. The phase comparators C compare the phase of clocks in the fast path and in the delayed path, and select a clock path where the phase difference is 360°. But the delay of delay unit changes over PVT variations. Therefore, the clock path differs even at a fixed clock frequency. To eliminate this effect of PVT variations, the delay units for additional delay are uniformly distributed over the delay chains and the CAC controls automatically the number of delay units to maintain the amount of additional delay equal to 90°. In the following, we show how the delay units for 90° phase shift are distributed. If x is the number of delay units for 0° phase shift, the total forward delay can be obtained by (1). If y is the number of additional delay units for the phase shift of T_{clk}/n , the total forward delay becomes (2). Let us assume that one delay unit is added at every m units in both paths. Then the total additional delay can be obtained as in (3) by dividing the total forward delay by m . This is equal to T_{clk}/n and yD . Finally, (4) can be obtained by solving (1), (2), and (3). With $n = 8$ and $m = 9$, one can generate a clock phase shifted by 90°. This is equivalent of adding one delay unit at every nine delay units for 90° phase shift. Furthermore, m and n are independent from the number of delay units for 0° and 90° phase shift. Therefore, PVT variations changing the x and y do not affect the phase of the output clock because the change in x and y are compensated by the change in unit delay, resulting in a phase shift of 90°.

$$Td + xD = T_{clk} \quad (1)$$

$$Td + (x + y)D = T_{clk} + \frac{T_{clk}}{n} \quad (2)$$

$$\frac{Td}{m} + \frac{(x + y)D}{m} = \frac{T_{clk}}{n} = yD \quad (3)$$

$$\Rightarrow yD \frac{(m - 1)}{m} = \frac{(Td + xD)}{m} = \frac{T_{clk}}{m}$$

$$\Rightarrow yD = \frac{T_{clk}}{(m - 1)} = \frac{T_{clk}}{n}$$

$$\Rightarrow m = n + 1 (n = 8, m = 9 \text{ for } 90^\circ \text{ phase shift}) \quad (4)$$

The jitter of the clock shifted by 0° is ± 13 ps and that of the clock shifted by 90° is ± 40 ps. In addition, the amount of phase

Fig. 10. Test waveforms of the proposed clock adjustment circuit.

shift can be controlled by JTAG_CONT to trim the setup/hold time and the data valid window. The covered range of the phase shift by JTAG_CONT is $\pm 20^\circ$ with ± 7 steps. Fig. 10 shows the test result of the proposed CAC. Two clocks phase shifted by 90° from the rising and falling edge of external clock are measured.

V. ON-CHIP TERMINATION

A. Termination Scheme

Termination has been used in high-data-rate systems to prevent unwanted reflections and improve signal integrity [9]. Off-chip termination has been widely used, but there is a difference between on-chip termination and off-chip termination. Off-chip termination has an unterminated stub composed of package parasitic and internal circuitry. This unterminated stub causes relatively large reflections compared with on-chip termination due to the impedance mismatch. Therefore, on-chip termination is developed to remove the effect of the unterminated stub and improve signal integrity. An on-chip input termination of the center-tapped-termination (CTT) type is designed by using CMOS transistors. Fig. 11(a) shows the input termination scheme of data pad. The off-chip driver (OCD) is activated during the read operation and the terminator is activated during the write operation. The impedance of the OCD and the terminator is controlled by digital codes generated by two PICs with reference resistors RQ and RT [10]. In 72-Mb DDR3 SRAM, RQ is 125 Ω and RT is 150 Ω . This means that the output impedance of the off-chip drivers is 25 Ω , which is equal to the board characteristic impedance, and the input impedance of terminators in the data pads is 75 Ω . Termination impedance is not matched to the board characteristic impedance to reduce the dc current dissipated in the terminator. To reduce input capacitance, 1/5 of the OCD transistors are used as terminators during nonread operation. Fig. 11(b) and (c) shows the simplified schematic of the terminator for the data pad and the address/control/clock pad. The terminator for the data pads consists of transistor arrays of nMOS and diode-connected nMOS pairs for pulldown, and pMOS and diode-connected pMOS pairs for pullup. The terminator for address/control/clock pads is composed of transmission gate arrays of nMOS and pMOS pairs. N -bit impedance codes for pullup and pulldown are generated from the PIC independently.

Fig. 11. (a) Input termination scheme of data pad. (b) Simplified schematic of terminator for data pad. (c) Simplified schematic of terminator for address pad.

Therefore, the pullup impedance code (Pn) and pulldown impedance code (Nn) are not always complementary.

B. Programmable Impedance Controller

The PIC generates digital impedance codes which make the output impedance of driver and the input impedance of the terminator close to R_Q and R_T within PIC resolution. Because of the different characteristics of nMOS and pMOS after fabrication, the PIC generates impedance codes for pullup and pulldown independently [10]. In the previously published approaches, the pulldown code is generated from the pullup code which has a quantization error [10]. As a result, the accuracy of the pulldown code is dependent upon the pullup code. To eliminate the dependency of the pulldown code on the pullup code, a new impedance code generation scheme is proposed. Fig. 12 shows the block diagram of the PIC. By feedback operation of AMP1 and a pMOS (M0), V_{ZQ} becomes V_{REF} ($= V_{DDQ}/2$). The reference current of V_{REF}/R_Q flows through M0 and R_Q . M3 copies the current in M0 to make pulldown impedance code and M1, M2, M4, and AMP2 copy the current in M0 to make pullup impedance code. Therefore, pulldown impedance code is independent from the pullup code. Control blocks control the size of the detector (NDET, PDET) to make the detector output (DCUR, UCUR) equal to V_{REF} . When DCUR and UCUR become V_{REF} , the bias condition of NDET and PDET is equal to that of R_Q and M0, which means that impedance of NDET and PDET is R_Q . But, due to the digital control of NDET and PDET, DCUR and UCUR

Fig. 12. Block diagram of the PIC.

have quantization error. Quantization error can be reduced by increasing the number of control bits and the resolution. The quantization error of the designed PIC is within 2% with five bits when R_Q is 125 Ω . This means that an error of 0.5 Ω exists in the off-chip driver when the target impedance is 25 Ω .

C. Linearity of Terminator

The input impedance of terminator should be linear to maintain equal channel environment for changing pad voltage and to improve the signal integrity and the input data valid windows in a system. Fig. 13(a) shows the linearity of an ideal terminator. In the ideal case, the relationship between pad voltage and input current is perfectly linear. But due to the nonlinear characteristics in transistors, there exists linearity error. Fig. 13(b) shows the linearity error of the terminator. The impedance of terminator is evaluated either by forcing a voltage to a pad and measuring the current flowing into a pad, or by measuring pullup impedance and pulldown impedance, respectively. Measuring pullup impedance and pulldown impedance, respectively, is just executed in test mode. Fig. 13(b) shows the result of forcing a voltage and measuring the current. The total linearity error of the terminator is $\pm 4.1\%$ over PVT variations. The linearity error is measured between 0.3 and 1.2 V.

D. Eye Diagram

Fig. 14 shows the eye diagram of the input data at a data rate of 1.5 Gb/s and a power supply of 1.5 V. In the case of no termination, the signal swing is larger than that in on-chip termination. But the noise and the reflections increase jitter and reduce the input data valid window. In the case of on-chip termination, the signal swing is reduced because of the dc current path in termination. But due to the reduced noise and reflections, the wider input data valid window is obtained. The data input valid window with on-chip termination is 480 ps at 750 mV ± 200 mV when the terminator impedance and the board characteristic impedance is 75 and 25 Ω , respectively. Included in the results are 10% PVT variations, 10% termination impedance variations, and system models.

Fig. 13. (a) Relationship between pad voltage and input current in ideal termination. (b) Linearity error of the designed terminator over PVT variations.

Fig. 14. Eye diagram of on-chip termination versus off-chip termination.

VI. HARDWARE RESULTS

Fig. 15 shows the hardware results at a data rate of 1.5 Gb/s and 750-MHz core frequencies. K is the external clock of 750 MHz. Data (DQ0, DQ1) of 1.5 Gb/s is well aligned with the echo clock (CQ, CQb). DQ latency is 2.1 ns from the time that an address is captured. Fig. 16 shows the chip micrograph of the 72-Mb SRAM.

VII. CONCLUSION

A 1.2-V 1.5-Gb/s 72-Mb DDR3 SRAM with the highest density and the smallest cell size to date in 6T SRAM is developed. The SMDL scheme is adopted to reduce main data line precharging power dissipation and the number of main data

Fig. 15. Hardware results.

Fig. 16. Chip micrograph of 72-Mb SRAM.

lines by half. A multiphase clock adjustment circuit is implemented to generate a clock phase shifted by 90° . With this clock adjustment circuit, input data can be sampled with optimized setup/hold window. Finally, on-chip termination controlled by PIC is developed to improve signal integrity and to remove the effect of the unterminated stub.

The SRAM is fabricated in a $0.10\text{-}\mu\text{m}$ CMOS process technology with five metals. The standby current in the memory cell and the power dissipation in the critical read/write path are minimized with the regulated voltage of 1.2 V. Off-current of the 72-Mb DDR3 SRAM measured is 80 mA. Gate oxide thicker than that of the memory cell is used to support the 1.5-V HSTL interface. Average core power dissipation with 50% read operation and 50% write operation is 1.2 W including off-current at 750 MHz and 1.5 V. Cell size and chip size are $0.845\text{ }\mu\text{m}^2$ and

TABLE I
FEATURES OF 72-Mb DDR3 SRAM

Organization	2Mx36, 4Mx18
Cycle time	750MHz
Data rate	1.5Gbps/pin
Power supply	1.5V ~ 2.5V
Operating voltage	1.2V(Regulated)
I/O interface	1.5V HSTL
Average core power	1.2W at 750MHz, 1.5V
Technology	0.1 μm CMOS
Tox	1.5 nm, 3.5 nm
Cell size	0.845 μm^2
Chip size	18865 μm x 8010 μm

151.1 mm^2 respectively. Table I summarizes the features of the 72-Mb DDR3 SRAM.

REFERENCES

- [1] S. Huang *et al.*, "High performance 50 nm CMOS devices for microprocessor and embedded processor core application," in *IEDM Tech. Dig.*, 2001, pp. 237–240.
- [2] S. Parihar *et al.*, "A high density 0.10 μm CMOS technology using low K dielectric and copper interconnect," in *IEDM Tech. Dig.*, 2001, pp. 249–252.
- [3] S. B. Kim *et al.*, "A 1.29 μm^2 full CMOS ultra-low power SRAM cell with 0.12 μm spacer-on-stopper (SOS) CMOS technology," in *IEDM Tech. Dig.*, 2001, pp. 253–256.
- [4] K. Tomita *et al.*, "Sub- μm^2 high density embedded SRAM technologies for 100 nm generation SOC and beyond," in *Symp. VLSI Tech. Dig. Tech. Papers*, 2002, pp. 14–15.
- [5] T. Chappell *et al.*, "A 2-ns cycle, 3.8-ns access 512-kb CMOS ECL SRAM with a fully pipelined architecture," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1577–1584, Nov. 1991.
- [6] H. Pilo *et al.*, "An 833 MHz 1.5 W 18 Mb CMOS SRAM with 1.67Gb/s/pin," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2000, pp. 266–267.
- [7] H.-C. Park *et al.*, "A 833 Mb/s 2.5V 4 Mb double data rate SRAM," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, vol. 464, Feb. 1998, pp. 356–357.
- [8] T. Saeki, "A 2.5 ns clock access, 250 MHz 256 Mb SDRAM with synchronous mirror delay," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1656–1668, Nov. 1996.
- [9] T. J. Gabara and D. W. Thompson, "A 200 MHz 100K ECL output buffer for CMOS ASICs," in *Proc. IEEE ASIC Seminar and Exhibit*, Sept. 1990, pp. P8/5.1–P8/5.4.
- [10] T. J. Gabara *et al.*, "Digitally adjustable resistors in CMOS for high-performance applications," *IEEE J. Solid-State Circuits*, vol. 27, pp. 1176–1185, Aug. 1992.

Uk-Rae Cho was born in Sang-Ju, Korea, in 1962. He received the B.S. degree in electronic engineering from Kyung-pook National University, Taegu, Korea, in 1985.

In 1984, he joined the Device Solution Network Division, Samsung Electronics Company, Yong-in, Korea, where he is currently a Project Leader of the SRAM Development Team. He holds eight international patents with 18 patents pending. His research interests include core circuits of ultrahigh-speed SRAM, high-bandwidth interface design, design for

test, device modeling, and analysis of BGA package substrate.

Tae-Hyoung Kim was born in Cheongju, Korea, in 1973. He received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, Korea, in 1999 and 2001, respectively.

In 2001, he joined the Device Solution Network Division, Samsung Electronics Company, Yong-in, Korea. Since 2001, he has been working on the design of high-speed SRAM memories. His research interests are analog circuits and high-speed I/O interface.

Yong-Jin Yoon was born in Seoul, Korea, in 1964. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University in 1987 and 1989, respectively. Since 1998, he has been working towards the Ph.D. degree at the same university.

In 1989, he joined the Device Solution Network Division, Samsung Electronics Company, Yong-in, Korea, where he is currently a Member of Technical Staff of the SRAM Development Team. His current research is in gh-speed synchronous SRAM.

Jong-Cheol Lee received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, Korea, in 1996 and 1998, respectively.

He joined Samsung Electronics Company, Yong-in, Korea, in 1998 and has been working on design and test of high-speed SRAM. His work experience includes design and analysis of chip critical timing path including decoder and sense amplifier, power analysis with voltage regulators, and floorplanning of high-speed SRAM. He is currently involved in chip verification for Samsung

DDR3 SRAMs.

Dae-Gi Bae was born on June 1, 1970, in Young-Ju, Korea. He received the B.S. degree in electric engineering from In-ha University, Korea, in 1996.

He joined the Memory Division, Samsung Electronics Corporation, Kiheung, Korea, in 1996, where he was involved in the circuit design of SRAM. From 1996 to the present, he has been working on the circuit design of high-speed synchronous SRAM memories.

Nam-Seog Kim was born in Seoul, Korea, in 1974. He received the B.S. degree in electrical engineering from Korea University, Seoul, Korea, in 1997 and the M.S. degree in electrical engineering from Seoul National University in 1999.

Since 1999, he has been a Member of Technical Staff with Samsung Electronics, Kiheung, Korea, where he is working on SRAM design and high-speed I/O. His research interests include low-power and high-performance circuits, clock recovery circuits, and high-speed link design.

Kang-Young Kim was born on June 6, 1970, in Kang-wondo, Korea. He graduated from Suwon Science College, Korea, in 1994.

He joined Samsung Electronics Corporation, Kiheung, Korea, in 1988, where he was involved in the circuit design of high-density NAND flash memories and MROMs. He has been working on the circuit design of high-speed SRAMs.

Kwang-Jin Lee was born on May 1, 1970, in Chonnam, Korea. He received the B.S. and M.S. degrees in electronics engineering from Korea University, Seoul, Korea, in 1994 and 1996, respectively. He is currently working toward the Ph.D. degree in electronics engineering at the same university.

He has been with Samsung Electronics, Kiheung, Korea, since 1996. His research interests are in high-speed memory design, special memory design such as CAM, and arm-based MCU design.

Young-Jae Son was born on November 17, 1971, in Busan, Korea. He received the B.E. degree in electronics engineering from Hanyang University, Korea, in 1994.

He joined the Memory Division, Samsung Electronics, Kiheung, Korea, in 1998, where he was involved in the circuit design of asynchronous fast SRAMs. From 1999 to the present, he has been working on the circuit design of ultrahigh-speed SRAMs.

Tae-Gyoung Kang was born on October 30, 1967, in Jeju, Korea.

He joined Samsung Electronics Corporation, Kiheung, Korea, in 1991, where he has been working on the layout of BiCMOS, NAND flash, 1T SRAM (DRAM cell SRAM interface), and high-speed SRAM.

Jeong-Suk Yang was born in Taejon, Korea, on March 27, 1976. She received the B.S. and M.S. degrees in electronic engineering from Chungnam University, Taejon, in 1999 and 2001, respectively.

She joined Samsung Electronics Company, Kiheung, Korea, in 2001, where she has been working on the circuit design of high-speed SRAM. Currently, she is involved in developing the next generation of high-speed SRAM.

Su-Chul Kim was born on Feb 23, 1970, in Pusan, Korea. He received the B.S. and M.S. degrees in electronic engineering from Korea University, Seoul, Korea, in 1995 and 2002, respectively.

In 1995, he joined the Samsung Electronics Company, Kiheung, Korea. Since then, he has been working on the design of 4M DDR and 8M SP and 32M DDR3 and high-speed SRAM memories. Currently, he is involved in developing 32M DDR SRAM memory.

Kwon-Il Sohn joined the Memory Division, Samsung Electronics Corporation, Kiheung, Korea, in 1999. From 1999 to the present, he has been working on the circuit design of ultrahigh-speed SRAMs.

Sung-Tae Kim was born on July 21, 1974, in Kwang-Ju, Korea. He received the B.S. degree in electronics engineering from A-Ju University, Korea, in 2001.

He joined the Memory Division, Samsung Electronics Corporation, Kiheung, Korea, in 2001, where he has been working on the circuit design of high-speed SRAM.

Kee-Sik Ahn was born on September 26, 1964, in Seoul, Korea. He received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 1987.

He joined Samsung Electronics Corporation, Kiheung, Korea, in 1987, where he was involved in the modeling of BiCMOS device. From 1992 to the present, he has been working on the circuit design of high-speed synchronous SRAMs and asynchronous fast SRAMs.

In-Yeol Lee was born on May 12, 1971, in Kyung-Nam, Korea. He received the B.E. degree in electronics engineering from Kyungbuk National University, Korea, in 1998.

He joined Samsung Electronics Corporation, Kiheung, Korea, in 1998, where he has been working on the circuit design of asynchronous fast SRAMs, high-speed DDR SRAM, and ternary content-addressable memory.

Hyun-Geun Byun was born on October 17, 1957, in Kyungbook, Korea. He received the B.S. degree in electronic engineering from Kyungbook National University, Taegu, Korea, in 1983.

He joined the Memory Development Division, Samsung Electronics, Korea, in 1983, where he was engaged in the development of low-power SRAM and high-speed SRAM.