

Received August 18, 2020, accepted October 4, 2020, date of current version October 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030099

An Embedded Level-Shifting Dual-Rail SRAM for High-Speed and Low-Power Cache

TAE HYUN KIM¹, HANWOOL JEONG², JUHYUN PARK³, (Member, IEEE),
HOONKI KIM⁴, (Member, IEEE), TAEJOONG SONG⁴, (Member, IEEE),
AND SEONG-OOK JUNG¹, (Senior Member, IEEE)

¹School of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, South Korea

²Department of Electronic Engineering, Kwangwoon University, Seoul 01897, South Korea

³Volume Product Design Group, DRAM Design Division, SK Hynix Inc., Icheon 467-701, South Korea

⁴Foundry Division, Samsung Electronics, Gyeonggi 445-701, South Korea

Corresponding author: Seong-Ook Jung (sjung@yonsei.ac.kr)

This work was supported by Samsung Electronics Company, Ltd., Device Solution (Variation-aware SRAM design techniques for low-power applications, 2015–2020).

ABSTRACT An embedded level-shifting (ELS) dual-rail SRAM is proposed to enhance the availability of dual-rail SRAMs. Although dual-rail SRAM is a powerful solution for satisfying the increasing demand for low-power applications, the enormous performance degradation at low supply voltages cannot meet the high-performance cache requirement in recent computing systems. The requirement of many level shifters is another drawback of the dual-rail SRAM because it degrades the energy-savings. The proposed ELS dual-rail SRAM achieves energy-savings by using a low supply voltage to precharge bitlines while minimizing the performance overhead by appropriately assigning a high-supply voltage to critical circuit blocks with effective level-shifting circuits. The sense amplifier embeds a level-shifting operation, thereby operating with a high supply voltage for a fast sensing operation. The proposed dynamic output buffer resolves the potential static current problem and improves the read delay. The number of level shifters is reduced using a proposed write driver, which conducts level-shifting and write-driving simultaneously. The proposed ELS dual-rail SRAM achieves low-power operation with 71.4% power consumption compared to single-rail SRAM with 72% performance overhead in circuit-level simulation, while the previous hybrid dual-rail SRAM shows 67.8% energy consumption with 270% performance overhead. In architecture-level simulation using Gem5 simulator with SPEC2006 benchmarks, the system with the ELS dual-rail SRAM caches shows, on average, 29% performance improvement compared to that of the system with the hybrid dual-rail SRAM caches.

INDEX TERMS Dual-rail SRAM, static random access memory, cache, energy-savings, Gem5 simulator, low-power operation, level shifter, output buffer, performance degradation, sense amplifier, SPEC2006 benchmarks, write driver.

I. INTRODUCTION

Technology scaling leads to highly integrated circuits for portable and Internet of Things (IoT) devices [1]. Because a long-lasting battery lifetime is inevitable for these devices, low-power operation has become the mainstream of logic circuit design in the past few years. To achieve low-power operation, the processor needs to operate with a relatively low supply voltage (VDD) to significantly reduce the dynamic power dissipated in the circuits.

The associate editor coordinating the review of this manuscript and approving it for publication was Yue Zhang¹.

A static random access memory (SRAM) is used as cache memory between the processor and main memory (usually dynamic memory, DRAM) [2]. The SRAM and processor are implemented on the same chip, which means that they share the supply voltage. The energy consumption in the caches is emphasized in [3], [4], which is approximately 12% to 45% of the core energy consumption, depending on the application, according to [3]. Thus, to achieve a low-power operation of the entire chip, it is necessary to reduce energy consumption in not just the processor, but the caches as well. However, in contrast to the processor, the SRAM has a low-bounded supply voltage requirement to meet the target yield

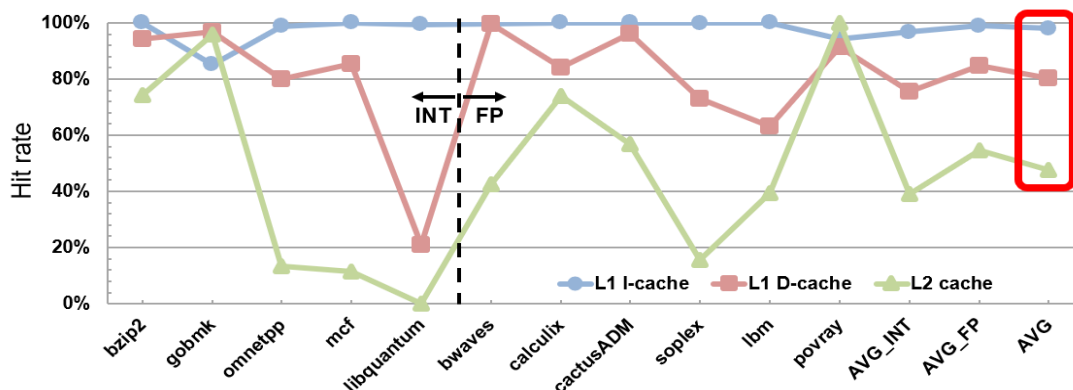


FIGURE 1. Hit-rate simulation results of L1 I/D caches and L2 cache.

[5]–[7]. This limitation has worsened in FinFET technology because the voltage difference between the supply voltage and threshold voltage (V_{th}) has decreased due to technology scaling [8].

A dual-rail architecture concept was suggested in [8], [10], and [11] for meeting the supply voltage requirement of SRAM. In the dual-rail architecture, the processor operates with a low supply voltage (VDDL), while the SRAM operates with a high supply voltage (VDDH). In this way, the total power consumed by the entire chip can be reduced while maintaining the target yield of the SRAM. Another solution is to adopt the dual-rail architecture inside the SRAM, which is called dual-rail SRAM [8]–[11]. The energy efficiency of the dual-rail SRAM can be optimized by appropriately distributing the VDDH and VDDL to the cell array, peripheral circuits, and control block in the SRAM. Several dual-rail SRAM schemes were introduced in [8]–[12].

Although energy consumption decreases with the use of dual-rail SRAMs, the performance overhead increases due to the peripheral circuits using VDDL.

The significant performance degradation, a critical disadvantage of the dual-rail SRAM, cannot meet the speed requirement of several applications [4] because more cache layers are used, and a high cache performance is required in recent computing systems, such as chip-multiprocessor (CMP) [13].

To achieve low-power operation in caches (SRAM) without the considerable performance overhead of conventional dual-rail SRAMs, we propose a new embedded level-shifting (ELS) dual-rail SRAM. The rest of this paper is organized as follows. Section II describes the motivation of this paper. Section III reviews previous dual-rail SRAMs. Section IV introduces the proposed ELS dual-rail SRAM and compares it with the conventional dual-rail SRAMs. Section V illustrates and analyzes the circuit-level simulation results. Section VI presents performance improvement in the proposed ELS dual-rail SRAM with architecture-level simulation. Finally, Section VII concludes the paper.

II. DESIGN ISSUES OF DUAL-RAIL SRAM FOR HIGH-SPEED LOW-POWER CACHE

This paper aims to achieve low-power operation in SRAM without a significant performance overhead.

First, the impact of cache performance on the computing system is simulated by observing the hit rates of L1 and L2 caches by using the architectural simulator Gem5 [14] with SPEC2006 benchmarks [15]. Some of these benchmarks are similar to each other, and thus a subset, which is a representative benchmark for the similar ones, can be selected [16]. Thus, five integer (INT) and six floating-point (FP) benchmarks are selected to be simulated. The cache configuration of the previous hybrid dual-rail SRAM [9], shown in Table 5, is used in the simulation; for other processor configuration parameters and the evaluation environment used in the simulation, please refer to Table 5 and Section VI. The hit-rate simulation results of L1 instruction/data caches (I/D-caches) and L2 cache are shown in Fig. 1, where the highlighted average values show the hit rates of L1 I/D-caches and L2 cache as 98%, 80%, and 48%, respectively. As the hit rates are high, the impact of cache performance on the computing system is confirmed to be substantial. Thus, a low-power operation in caches needs to be achieved with a reasonable performance degradation.

Next, characteristics of the previous dual-rail SRAMs are introduced to determine the important issues in designing dual-rail SRAM. Fig. 2 shows the sub-blocks and signal flow of a conventional single-rail SRAM synchronized with the clock. The data array has a size of 16 KB, comprising four 32 KB bitcell arrays, that is, 256 rows (wordlines, WLs) by 128 columns (bitlines, BLs). Two horizontally aligned bitcell arrays form a group and operate simultaneously [17]. The four-bitwise bit-interleaving is adopted with 4 to 1 read and write column multiplexers (MUXs). Thus, the input data (DIN) and output data (Q) are 64 bits in length each. Several signals enter the SRAM, including the clock (CLK), write enable (WEN), address (ADDR), and DIN, and Q is transferred out from the SRAM. In the read or write operation, 11-bit ADDR signals are used. Among these multibit

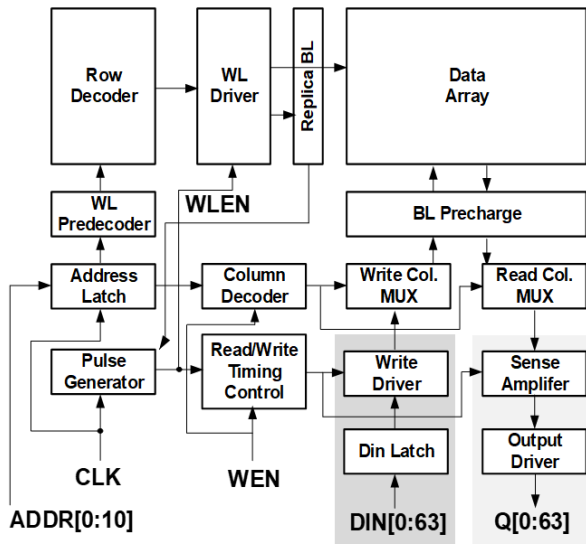


FIGURE 2. SRAM sub-blocks and signal flow.

ADDRs, one bit is used to select one of the two groups of bitcell arrays, eight bits are used to specify one WL selected from the 256 WLS, and two bits are used for the 4 to 1 MUX. WEN determines the read or write operation. In the dual-rail SRAM, how the VDDL and VDDH are distributed to the sub-blocks in Fig. 2 determines the SRAM operation as well as the position and number of level shifters, and thus, the performance and energy consumption.

The interface dual-rail SRAM [11] is the SRAM used in the dual-rail architecture mentioned in the previous section, and thus, all blocks use VDDH. At the processor–SRAM interface, several VDDL signals, such as CLK, WEN, ADDR, and DIN are conveyed to the SRAM, and hence level shifters are required at the interface. Because all the blocks use VDDH, energy saving within the SRAM is not expected for the interface dual-rail SRAM, and thus energy saving of the entire chip achieved using the dual-rail SRAM is limited. In contrast to the interface dual-rail SRAM, the array and hybrid dual-rail SRAMs use dual supply voltages (VDDH and VDDL) [8]–[10]. Because several blocks in these dual-rail SRAMs related to performance (read delay) use a lower supply voltage (VDDL), a considerable performance degradation occurs. Furthermore, in all the previous dual-rail SRAMs, a large number of level shifters required at the interfaces of the VDDH and VDDL domains consume large energy and occupy large area [18]. In addition, a dual-rail SRAM in which only the WL driver and data array use VDDH, while other blocks use VDDL is also suggested in [12]. However, this SRAM suffers from writability degradation due to the weak write driver.

Therefore, two important issues should be considered with respect to dual-rail SRAMs. First, the VDDL and VDDH domains need to be appropriately determined to improve the energy-delay product (EDP) by maximizing the energy-saving efficiency without significant performance degrada-

tion while satisfying the target yield. Second, an efficient level-shifting circuit is required to minimize the energy consumption. The proposed ELS dual-rail SRAM resolves these issues, and a detailed analysis is presented in the following sections.

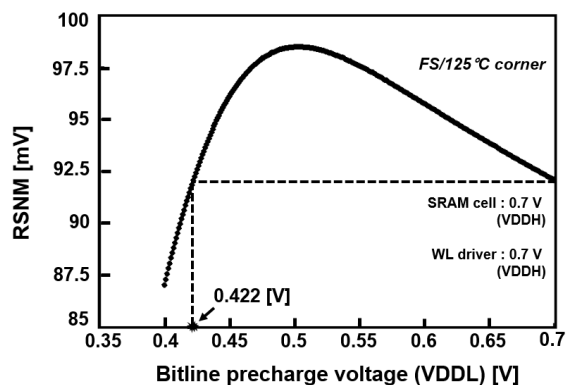
III. PREVIOUS DUAL-RAIL SRAMs

A. ANALYSIS AND COMPARISON OF PREVIOUS DUAL-RAIL SRAMs

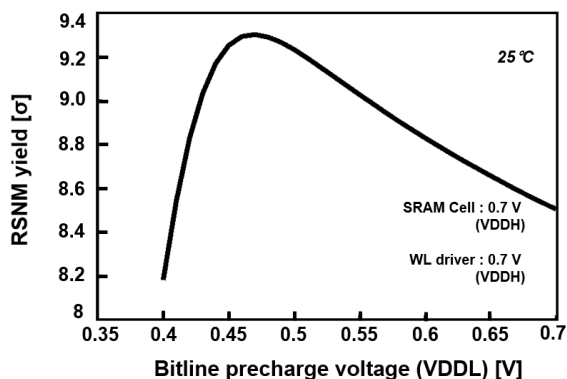
In this section, a detailed analysis and comparison of the previous dual-rail SRAMs are conducted.

First previous dual-rail SRAM is the interface dual-rail SRAM in which all SRAM blocks operate with VDDH. As pull-up PMOS is not completely turned off when the VDDL signal drives a gate using VDDH [18], a level shifter is required for all input signals and it cause energy, performance and area overheads. Especially, as mentioned in Section II, energy saving of the entire chip is limited because the interface dual-rail SRAM operates only with VDDH, same as the single-rail SRAM.

Next dual-rail SRAM is the array dual-rail SRAM. The motivation of this dual-rail SRAM lies in the fact that VDDH is mainly required for the bitcell to ensure the yield. In other words, only the blocks related to the bitcell, including the bitcell array itself, must be in the VDDH domain. Most of the other blocks, including the BL precharge circuit, are in the VDDL domain. Because the energy consumption in the SRAM is dominated by BL precharging, it can be considerably reduced using VDDL to precharge BL. However, with respect to performance, two potential critical paths, that is, the address path including the row decoder, and the WL enable (WLEN)-generating path are under the VDDL domain, and thus, performance degradation is inevitable in the array dual-rail SRAM. Furthermore, the read-out path, including the sense amplifier, is also in the VDDL domain, which aggravates the performance degradation. The level shifters are required to control the WL drivers with VDDH because the control block is in the VDDL domain, while the WL driver is in the VDDH domain. Although energy can be saved by precharging BL with VDDL, the VDDL level should be carefully selected because it affects the read static noise margin (RSNM) of the SRAM [8]. Fig. 3(a) shows the RSNM of a conventional 6T SRAM cell at the worst-case corner (fast NMOS, slow PMOS, and 125°C). As the BL precharge voltage decreases, the RSNM of the SRAM with a cell VDD of VDDH initially increases and then decreases. The initial increase in the RSNM with decreasing BL precharge voltage is due to the decrease in the read disturbance current at the node storing “0” in the SRAM cell. However, as the BL precharge voltage further decreases, the RSNM decreases because the node storing “1” in the SRAM cell is discharged more, leading to more data flip in the SRAM cell. The RSNM yield shows a similar tendency in Fig. 3(b). Not only the read stability (RSNM) but also the write ability should be considered. Fig. 4 shows the write margin in the worst-case corner (slow NMOS, fast PMOS, and –40°C) when the



(a)



(b)

FIGURE 3. (a) RSNM of the 6T SRAM at the worst-case corner and (b) RSNM yield change depending on the bitline precharge voltage.

write-voltage wordline (WVWL) is used as a write ability metric [19], [20]. WVWL is a suitable metric for comparing the write margin when the strength of the write driver varies. In general, the write operation depends on the NMOS in the write driver, which drives “0” to the SRAM cell. As the voltage level of input data to the write driver decreases, the strength of the NMOS is reduced, leading to a decrease in the WVWL. Thus, the input data for the write driver in the array dual-rail SRAM should be up-shifted to the VDDH level to achieve a strong drivability of the NMOS for attaining a sufficient write margin. Thus, the level shifters are essential for the input data. Because the data write-in path includes the write driver in the VDDH domain and the read-out paths include a sense amplifier in the VDDL domain, the read and write column MUXs are also separated as the VDDL and VDDH domains, respectively.

Last dual-rail SRAM is the hybrid dual-rail SRAM, which also reduces the energy consumption by using VDDL to precharge the BL, similar to the array dual-rail SRAM. The read-out path, including the sense amplifier and output buffer, are also in the VDDL domain. On the other hand, the blocks including the control block, row decoder, and WL driver are in the VDDH domain. In this way, the critical path, except for the data read-out path, is in the VDDH domain. Thus, the

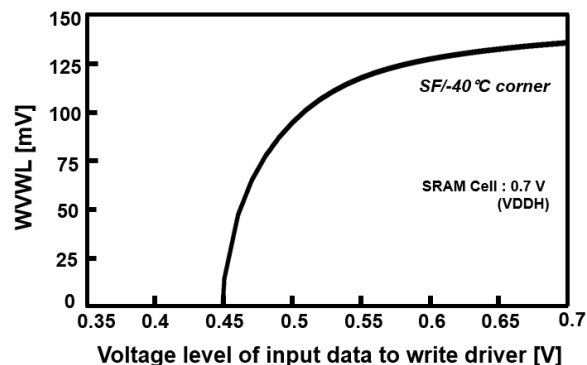


FIGURE 4. Change in the write voltage wordline (WVWL) margin according to the voltage level of the input data for the write driver at the write worst-case corner (SF/-40°C).

disadvantage of performance degradation in the array dual-rail SRAM can be mitigated. However, a large number of level shifters are still required for the input signals. Although the control block is in the VDDH domain, some control signals are transferred to the peripheral circuits with VDDL because the circuits operate with VDDL [9].

The supply voltage-domain distribution of these dual-rail SRAMs is summarized in Table 1 at the end of Section IV, and the overall comparison for characteristics of the three previous dual-rail SRAMs is summarized in Table 4 at the end of Section V.

B. LIMITATION OF THE HYBRID DUAL-RAIL SRAM

Although the hybrid dual-rail SRAM has theoretical advantages over the interface and array dual-rail SRAMs, it has some limitations.

First, placing the most read-out path in the VDDL domain leads to an enormous increase in EDP, because it significantly degrades the performance, even with large energy savings in the BL precharging. As VDDL decreases, this problem worsens. In general, BL development delay is dominant in read delay. However, as VDDL decreases, the sense amplifier delay occupies a larger portion of the total read delay. The reason for this is as follows. The sense amplifier delay is significantly affected by the VDDL level because the sense amplifier operates with VDDL, while the BL development delay is not much affected because WL and SRAM cells, which mainly affect the BL development, operate with VDDH.

Second, a large number of level shifters are still placed at the SRAM–processor interface. Most of the level shifters are required for the input data. In the level shifter, a large NMOS is required to consider the contention between the NMOS with VDDL and PMOS with VDDH, which also causes a large static current. Thus, they cause noticeable energy consumption and area overhead.

Finally, using VDDL to transfer some control signals causes a performance overhead. This affects not only the performance but also the energy consumption because it determines the WL pulse width and then the BL development

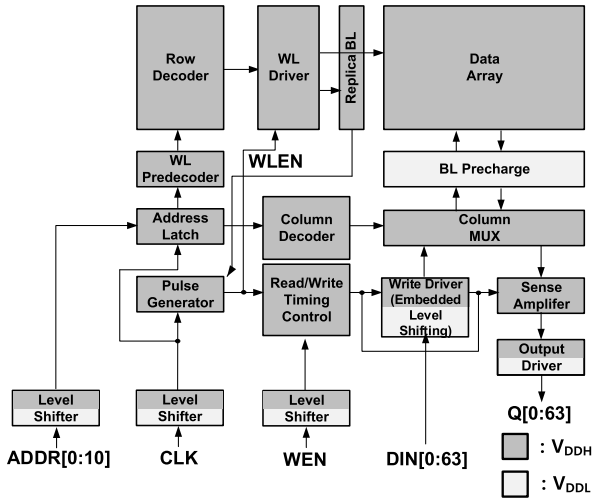


FIGURE 5. Proposed dual-rail SRAM sub-blocks.

time. Furthermore, the timing issue between the control signals must also be carefully considered because the VDDH and VDDL control signals are used together.

IV. PROPOSED ELS DUAL-RAIL SRAM

In this section, an embedded level-shifting (ELS) dual-rail SRAM is proposed to resolve the limitations mentioned in Section III.B. The proposed ELS dual-rail SRAM aims to appropriately distribute the supply voltage domain (VDDH or VDDL) to each sub-block and achieve large energy savings with a reasonable performance. For this purpose, the following techniques are proposed.

First, the structure of the column MUX is optimized to efficiently convey the read and write data to the BL by using a transmission gate. Second, all control signals for the peripheral circuits, including the write driver and sense amplifier, are transferred with VDDH. In this way, the performance degradation in the control block is minimized. Third, the performance degradation in the sense amplifier is substantially improved using VDDH with a negligible energy overhead. Additionally, a dynamic output buffer is proposed that cuts the potential static current path in the buffer while enhancing the performance. Finally, the number of level shifters for the input data is reduced by embedding a level-shifting operation in the proposed power-gated cross-coupled inverter-based level-shifting write driver. Thus, the energy consumption for level shifting decreases with the proposed write driver, which comprises a power-gated cross-coupled inverter-based level-shifting circuit and a clock-gated driver.

A. BLOCK DIAGRAM

Fig. 5 shows the block diagram of the proposed ELS dual-rail SRAM where the SRAM cells and WL driver are in the VDDH domain, and the BL precharge block is in the VDDL domain. The control block and row decoder are in the VDDH domain. The read and write MUXs are integrated into the column MUX, which comprises transmission gates (TGs)

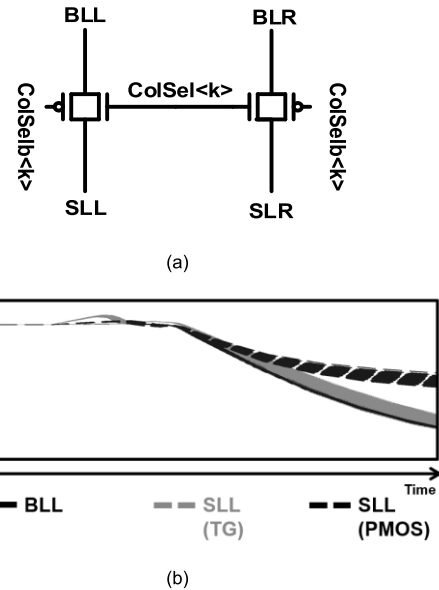
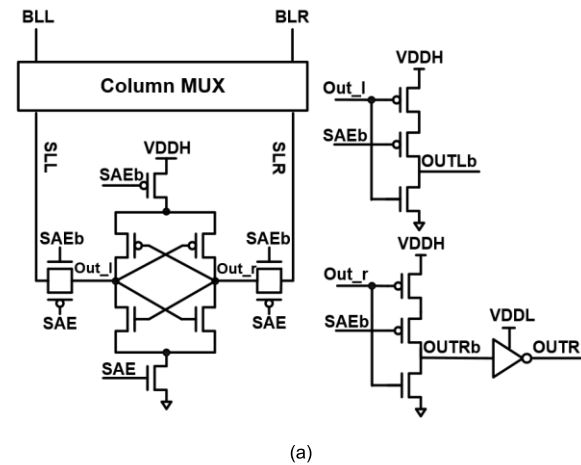


FIGURE 6. (a) Structure of column MUX and (b) waveform for bitline and sensing line discharge operation of column MUX with TG and conventional PMOS.

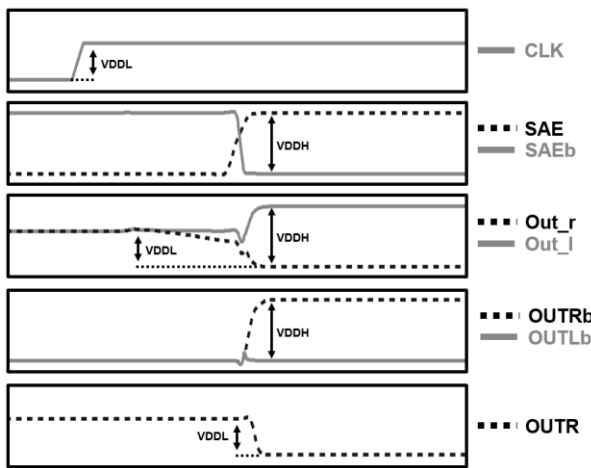
[21]. In Fig. 6(a), the column MUX operates with two VDDH control signals, ColSel and ColSelb. SLL and SLR stand for the sensing line pair nodes between the sense amplifier and column MUX. To explain the advantage of the column MUX, Fig. 6(b) shows the discharging operation in BLL and SLL for both the TG column MUX and PMOS read MUX. The TG column MUX completely conveys the BLL discharge to the SLL, while the PMOS column MUX does not because the PMOS is limited in driving the BL below the gate to its source voltage ($|V_{gs}|$). Thus, the performance degradation can be mitigated using the TG column MUX. Furthermore, the “1” data can be successfully transmitted from the SL to the BL through the column MUX in the write operation because the column MUX is shared in both the read and write operations, as shown in Fig. 5.

B. CONTROL SIGNALS

All control signals in the proposed ELS dual-rail SRAM are generated and transferred with VDDH, with several advantages. The BL precharge control signal with VDDH has a fast-rising transition delay compared to that with VDDL, and the column MUX control signal (ColSel) with VDDH helps the column MUX to convey the BL discharge much faster, leading to a decrease in the read delay. Moreover, the rising delay of the sense amplifier enable (SAE) with VDDH is much faster than that of the SAE with VDDL. Because the speed of WL falling transition is affected by that of the SAE rising transition, a fast rising of SAE shortens the WL pulse width, which reduces the BL precharge energy by preventing unnecessary BL discharge. In the write operation, ColSel with VDDH also enhances the strength for writing “0” data, thus enhancing the write ability.



(a)



(b)

FIGURE 7. (a) Level-shifting sense amplifier (DSTA-VLSA) [13] and output buffer design for the proposed dual-rail SRAM and (b) circuit operation.

Additionally, the design complexity of the control block is reduced when only control signals with VDDH exist because the timing issue between the control signals with VDDL and those with VDDH does not occur.

C. LEVEL-SHIFTING SENSE AMPLIFIER AND OUTPUT BUFFER

The level-shifting sense amplifier embeds the level-shifting operation by using a voltage latch-type sense amplifier with double switches and transmission gate access transistors (DSTA-VLSA) [22]. Fig. 7(a) and (b) show the level-shifting sense amplifier (DSTA-VLSA) with a dynamic output buffer circuit and the operational waveforms of the sense amplifier in the read operation, respectively.

The level-shifting sense amplifier includes the head switch to cut off the DC static path from VDDH to VDDL because the sense amplifier operates in the VDDH domain, while the BL precharge level is VDDL. The pass gate comprises TGs for successful discharge at the out_l and out_r nodes shown in Fig. 7(a). The level-shifting sense amplifier operates as

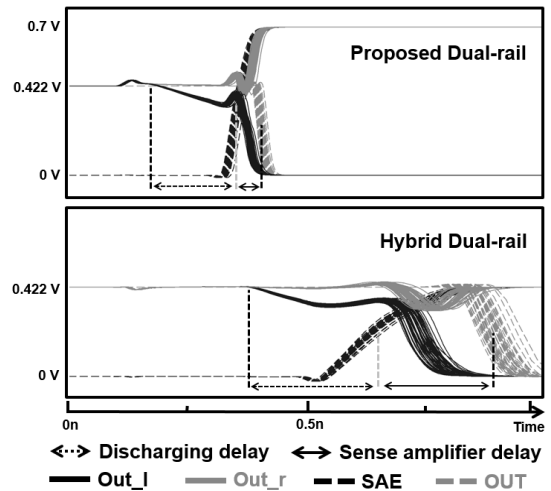


FIGURE 8. Waveforms for sensing operation of the proposed and hybrid dual-rail SRAMs.

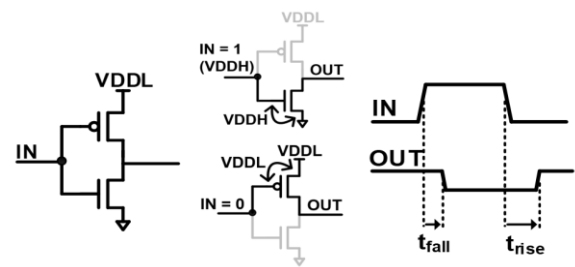


FIGURE 9. Problem that can occur when the VDDH signal drives the VDDL gate. Inverter gate, V_{gs} , and $|V_{gs}|$ for NMOS and PMOS and input and output signals from left to right.

follows. First, the out_l and out_r nodes start to develop as the BLs and SLs are developed by an SRAM cell. When the voltage difference between out_l and out_r becomes large enough, SAE activates the sense amplifier. After the head and foot switches are turned on by the SAEb and SAE, the voltage difference is amplified to VDDH by cross-coupled inverters. Then, the output data (OUTR) is transferred out through the output buffer. The details about timing of SAE that should be carefully considered are explained in Section V.

Fig. 8 shows the waveforms related to the sensing delay of the proposed and hybrid dual-rail SRAMs. As mentioned in Section III, as VDDL decreases, the delay of the conventional sense amplifier increases and occupies an enormous portion of the total read delay. In the proposed ELS dual-rail SRAM, the sense amplifier delay significantly decreases using VDDH.

In Fig. 7(a), the dynamic output buffer comprises a dynamic inverter controlled by the SAEb and a secondary inverter. Both OUTLb and OUTRb are precharged to VSS during the BL precharge phase because out_l and out_r are precharged to VDDL. The pull-up path is cut off as SAEb holds VDDH. Thus, similar to a dynamic logic gate, only a VSS to VDDH rising transition is possible at the OUTLb or OUTRb nodes, leading to only falling transition from VDDL to VSS at the output of the secondary inverter (OUTR). Fig. 9 explains the performance advantage of this

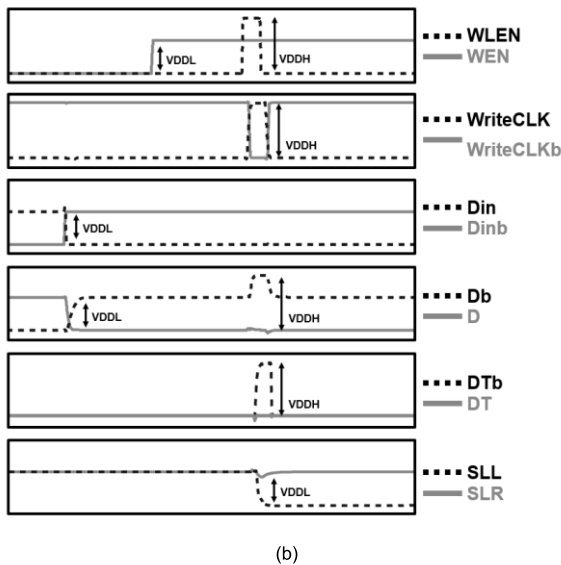
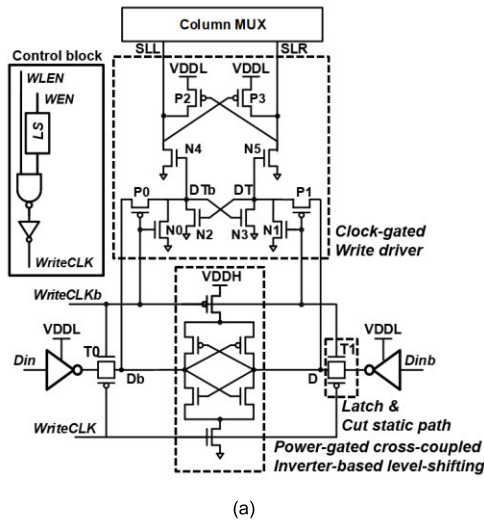


FIGURE 10. (a) Proposed power-gated cross-coupled inverter-based level-shifting write driver circuit and (b) operation of the proposed write driver.

falling transition. When the VDDH signal drives a gate using VDDL, the V_{gs} of NMOS becomes VDDH according to the rising transition of IN. On the other hand, $|V_{gs}|$ of PMOS becomes VDDL according to the falling transition of IN. This means that the falling transition of OUT is much faster than the rising transition. Thus, the delay of the dynamic output buffer can be minimized using only fast falling transition in the OTR node. In addition, the potential static current path caused by the VDDL precharge voltage is cut because SAEB with VDDH in the precharge phase turns off PMOS in the output buffer.

D. POWER-GATED CROSS-COUPLED INVERTER-BASED LEVEL-SHIFTING WRITE DRIVER

Fig. 10(a) shows the proposed power-gated cross-coupled inverter-based level-shifting write driver. The proposed write

driver simultaneously conducts level-shifting and write-driving. Fig. 10(b) shows the signal waveforms during the write operation. There are three input signals to the write driver. WriteCLK is the VDDH pulse signal generated with up-shifted WEN (VDDH) and WLEN, as shown in the left-top box in Fig. 10(a). Thus, it is activated with the WLEN only in the write operation. Din and Dinb are complementary input data with VDDL.

In the BL precharge phase, where both WEN and Write-CLK are “0,” the power-gated cross-coupled inverter-based level-shifting circuit, which comprises the power-gated cross-coupled inverters, is turned off, while the transmission gates T0 and T1 are turned on. Thus, the D and Db nodes in the write driver have the same values as Din and Dinb, respectively. This indicates that the changes in the input data are conveyed to D and Db during the BL precharge phase. Simultaneously, P0 and P1 are turned off, while N0 and N1 are turned on in the clock-gated write driver. Thus, N4 and N5 are turned off to not disturb the precharging operation in the SLL and SLR because the DT and DTb nodes hold the VSS.

Next, the write operation steps are as follows. After WEN is enabled, WriteCLK is activated at the rising edge of WLEN. Subsequently, T0 and T1 are turned off, and thus, the changes in the input data do not affect D and Db anymore. Simultaneously, the cross-coupled inverters start operating with the supplied VDDH and VSS. Because the voltage difference between D and Db is as much as VDDL, the cross-coupled inverters can easily pull up the VDDL node (D or Db) to VDDH, as shown in the waveform in Fig. 10(b). The level-shifted input data are conveyed to the DT and DTb nodes as P0 and P1 are turned on, while N0 and N1 are turned off. At this time, one of the cross-coupled NMOSs, that is, N2 or N3, makes DT or DTb hold “0” according to the input data. Then, N4 or N5 drives the SLL or SLR, and thus, BLL or BLR, through the TG column MUX with the strength of VDDH, leading to a large write margin, as explained in Section III.A.

The level-shifting based on cross-coupled inverters does not suffer from the serious contention current because the voltage difference in the complementary nodes, D and Db, can be easily amplified to VDDH without a large contention when the voltage difference of the input data is as large as VDDL. Thus, the large static current problem of the level shifter is resolved, and the energy consumed by level shifting is reduced. In the write operation, energy saving is maximized because most of the level shifters exist for the input data. In addition, T0 and T1, which play the roles of data latch, cut the possible static current path from VDDH to VDDL.

With the proposed write driver, 64 level shifters are removed, which corresponds to the width of input data. Because level shifting logic embedded in the proposed write driver consumes much less energy than level shifters, write energy consumption in the ELS dual-rail SRAM is reduced. Additionally, the area overhead of level shifters can be also mitigated because the level shifting circuit in proposed write

TABLE 1. Characteristics comparison of the proposed ELS dual-rail SRAM and previous dual-rail SRAMs.

Comparative Characteristics	Interface	Array	Hybrid	ELS (proposed)
Bitline precharge level	VDDH	VDDL	VDDL	VDDL
Charge control signal level	VDDH	VDDL	VDDL	VDDH
Read and write MUXs	Separated read and write MUXs	Separated read and write MUXs	Separated read and write MUXs	Shared TG column MUX
Sense amplifier (supply voltage)	VLSA (VDDH)	VLSA (VDDL)	VLSA (VDDL)	Level-shifting sense amplifier (VDDH)
Sense amplifier enable signal level	VDDH	VDDL	VDDL	VDDH
Output buffer (supply voltage)	Static output buffer (VDDH)	Static output buffer (VDDL)	Static output buffer (VDDL)	Dynamic output buffer (VDDH)
Write driver (supply voltage)	Single-voltage write driver (VDDH)	Single-voltage write driver (VDDL)	Single-voltage write driver (VDDL)	Power-gated cross-coupled inverter-based level-shifting write driver (VDDH/VDDL)
Input data level shifters	Needed	Needed	Needed	

TABLE 2. Impact of the characteristics of the proposed ELS dual-rail SRAM compared to the previous dual-rail SRAMs.

Characteristics of the ELS dual-rail SRAM	VS. Interface	VS. Array	VS. Hybrid
Bitline precharge with VDDL	Read delay ↑ Read/Write Energy ↓↓↓	No change	No change
Precharge control signal with VDDH	No change	Read delay ↓↓ Read/Write Energy ↑	Read delay ↓↓ Read/Write Energy ↑
Shared TG column MUX	No change	Read delay ↓	Read delay ↓
Level-shifting sense amplifier	Read delay ↑	Read delay ↓↓↓ Read Energy ↑	Read delay ↓↓↓ Read Energy ↑
Sense amplifier enable signal with VDDH	No change	Read delay ↓↓ Read/Write Energy ↓	Read delay ↓↓ Read/Write Energy ↓
Dynamic output buffer	No change	Read delay ↓	Read delay ↓
Power-gated cross-coupled inverter-based level-shifting write driver	Write energy ↓↓	Write energy ↓	Write energy ↓

The arrows stand for, ↑/↓ : slightly increase/decrease, ↑↑/↓↓ : normally increase/decrease, ↑↑↑/↓↓↓ : dramatically increase/decrease.

driver, for which small offset voltage is not required, doesn't need to be large as the level shifter.

The overall characteristic comparison of the proposed and previous dual-rail SRAMs is summarized in Table 1 and the impact of their characteristics is summarized in Table 2.

V. CIRCUIT-LEVEL SIMULATION RESULTS

Based on a quantitative comparison between the single-rail SRAM, previous dual-rail SRAMs, and proposed dual-rail SRAM using a circuit-level HSPICE simulator, the energy savings and performance advantage of the proposed ELS dual-rail SRAM can be verified. The 14-nm FinFET model based on parameters, characteristics, and variation information reported in [23] and [24] is used for the simulation. The key I-V characteristics of the model are listed in Table 3. The effect of variation sources in FinFET, including the work function variation and line edge roughness [25], can be modeled with the variation in V_{th} . Equation (1) shows Pelgrom's

model:

$$\sigma V_{th} = \frac{A_{vt}}{\sqrt{Length \times Width}}, \quad (1)$$

where A_{vt} is the Pelgrom coefficient, and its values for NMOS and PMOS are set as 0.757 and 0.87 mV · μm, respectively, according to [26].

The SRAM architecture shown in Fig. 1 is used for the simulations [17]. The bitcell arrays comprise 256 rows and 128 columns, and the four-bitwise bit-interleaving is applied. A high-density SRAM cell (1:1:1, PU:PD:PG) [23] is used, and the cell height and width are estimated as 0.14 μm and 0.356 μm, respectively. The parasitic capacitance for WL, BL, and all control signals, including SAE, ColSel, Write-CLK, WLEN, and address signals, is set as 0.27 fF/μm based on [27]. The input signals, CLK, WEN, address bits, and input data are applied to SRAM with VDDL level for the dual-rail SRAMs. In the single-rail SRAM, all SRAM sub-blocks are under the VDDH domain, and the input signals are also at VDDH level. The VDDH level is set as 0.7V, which is the

TABLE 3. I-V characteristics of the used 14-nm FinFET model.

Characteristics	NMOS	PMOS
Nominal supply voltage	0.7 V	
On current	0.85 mA/ μ m	0.72 mA/ μ m
Off current	1 nA/ μ m	1 nA/ μ m
DIBL	47.7 mV/V	50.8 mV/V
Subthreshold swing	72.5 mV/dec	83.5 mV/dec
Saturation threshold voltage	268 mV	-315 mV

The threshold voltage is estimated with the constant current method by measuring V_{gs} when the drain current per effective width is 10^{-5} A/ μ m in the saturation region.

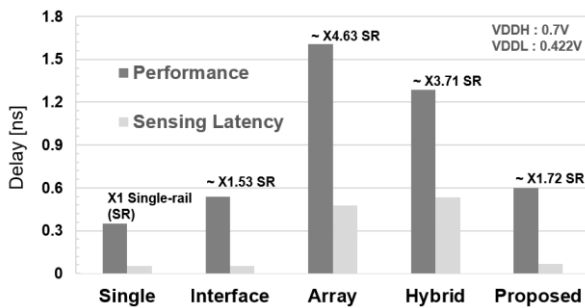


FIGURE 11. Performance and sensing latency for all SRAM schemes.

nominal supply voltage in the 14-nm FinFET model. For a fair comparison between the single-rail and dual-rail SRAMs, the RSNM condition should be same. Because the RSNM at 0.422V is the same as that at 0.7V (VDDH) corresponding to single-rail supply voltage, as shown in Fig. 3(a), the VDDL level is set as 0.422V.

The sigma of the offset voltage (σV_{OS}) is an important parameter for determining the SRAM performance and sensing yield. The target σV_{OS} is set as 20mV to meet the conventional industry design target [22]. The conventional voltage latch-type sense amplifier (VLSA) [22] with VDDH is used for the single-rail and interface dual-rail SRAM. The VLSA with VDDL is used for the array and hybrid dual-rail SRAM, and the level-shifting sense amplifier is used for the proposed dual-rail SRAM. Fin numbers 1 and 2 are used for pull-up and pull-down transistors, respectively, for all sense amplifier types to meet the target σV_{OS} .

The timing of the SAE signal is also important. In this simulation, SAE is determined based on the read access pass yield (RAPY), which considers both the BL voltage development and sense amplifier offset voltage variations [28]. The replica circuit is tuned to generate SAE signal that can meet target RAPY. The SAE signal timing is set separately for each SRAM scheme to meet the six-sigma RAPY.

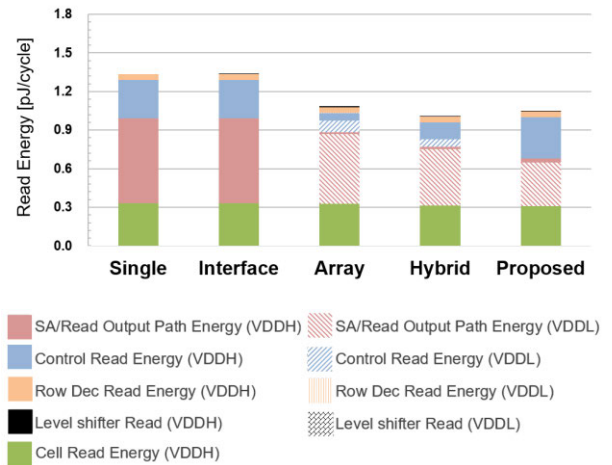


FIGURE 12. Energy consumption during the read operation for all SRAM schemes.

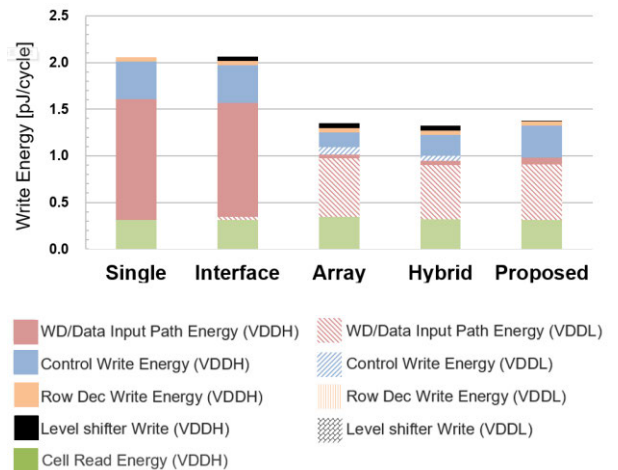


FIGURE 13. Energy consumption during the write operation for all SRAM schemes.

After the analysis for several level-shifter candidates [29]–[31], MWCMBH [29] is decided as a level-shifter design because of its low power consumption.

A. PERFORMANCE AND READ/WRITE ENERGY ANALYSES

The sensing latency and performance of each SRAM scheme when VDDL is 0.422V, is compared in Fig. 11. The performance is the read delay from CLK to the output data, and the sensing latency is the delay from SAE to the output data. The performance and sensing latency are the six-sigma worst-case delays estimated by applying the tail-fitting method [32] to 5000 Monte Carlo simulation results. The numbers on top of each performance graph indicate normalized performance to that of single-rail SRAM.

As mentioned in Section III, the interface dual-rail SRAM shows a delay most similar to that of the single-rail SRAM,

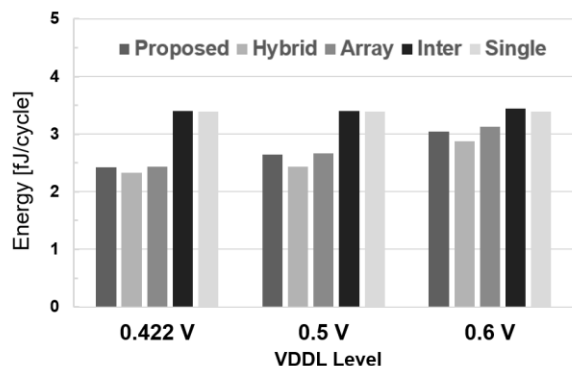


FIGURE 14. Energy consumption along various VDDL levels for all SRAMs.

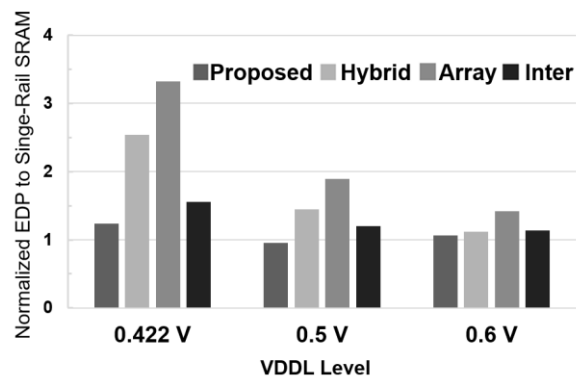


FIGURE 15. EDP normalized to the single-rail SRAM along various VDDL levels for dual-rail SRAMs.

except for the performance overhead of the level shifters at the interface. However, the array dual-rail SRAM has a lot of blocks in the VDDL domain, and thus, significant performance degradation is observed. The hybrid dual-rail SRAM also shows considerable performance degradation because the blocks in the read-out path, including the sense amplifier, operate slowly in the VDDL domain. In addition, the slow transition of the control signals also degrades the performance. On the other hand, the proposed ELS dual-rail SRAM shows relatively small performance degradation thanks to the level-shifting sense amplifier, dynamic output buffer, and fast transition control signals. As a result, the read delay of the proposed dual-rail SRAM is only 72% slower than that of the single-rail SRAM, whereas the read delay of the array and hybrid SRAMs is 4.6 and 3.7 times, respectively, that of the single-rail SRAM.

Fig. 12 and Fig. 13 compare the energy consumption during the read and write operations. The total energy consumption is divided into the energy consumed by each functional unit. A legend in the bottom shows each unit's name and the voltage domain. The cell read/write energy presents energy consumed in the cell. The SA/read output path energy covers the precharge energy after the read operation and energy consumed in the sense amplifier and output buffer. The WD/data input path energy covers the precharge energy after the write operation and energy consumed in the write driver. The control read/write energy presents energy consumed in the control block, including energy consumed to deliver control signals, and the row decoder read/write energy presents the energy consumed in the WL predecoder and WL driver. Finally, the level shifter energy presents the energy consumed by the level shifters.

The read energy consumption in the interface dual-rail SRAM is slightly larger than that in the single-rail SRAM owing to the presence of level shifters at the interface. However, the added energy consumption from the level shifters is not substantial because the level shifters for input data do not operate in the read operation. The read energy consumption of the array dual-rail is slightly larger than that of the proposed dual-rail SRAM despite many VDDL blocks. This is because

the WL pulse width is much longer than that in the proposed SRAM scheme, as the level shifter is placed on the path from the replica circuit (SAE-generating circuit) to the WL driver [3], leading to a larger BL development.

For the hybrid and proposed ELS dual-rail SRAMs, although the proposed dual-rail SRAM consumes more energy in the control block than the hybrid dual-rail SRAM owing to the many VDDH control signals, its BL precharge energy is smaller because of the relatively short WL pulse width, thanks to the fast transition of the SAE signal. Furthermore, the energy consumption of the read-out path in the VDDH domain retains a very small portion among the total energy consumption, as shown in Fig. 12, meaning that the additional energy consumption in the level-shifting sense amplifier is very small. As a result, the proposed dual-rail SRAM achieves enough read energy saving, even though it consumes slightly more energy than the hybrid dual-rail SRAM.

The write energy consumption difference between the single-rail and the interface dual-rail SRAMs is relatively larger than the read energy consumption because most level shifters contribute to the write operation for conveying the input data. The energy savings in the write operation of the array dual-rail SRAM are not notable because the write driver using up-shifted input data consumes similar energy to that of the other SRAMs. Additionally, the long WL pulse width also increases the write energy.

In the proposed ELS dual-rail SRAM, the proposed write driver removes the level shifters used for the input data by integrating level-shifting and write-driving functions. The energy consumption in the level shifters is clearly reduced in the proposed dual-rail SRAM compared to that in the interface, array, and hybrid dual-rail SRAMs. On the other hand, the energy consumption in the WD/data input path slightly increases. Thus, the summation of energy consumption in the level shifters and WD/data input path decreases thanks to the energy-efficient level-shifting of the proposed write driver. Thus, the proposed dual-rail SRAM shows similar write energy consumption as that of the hybrid dual-rail

TABLE 4. Comprehensive comparison between the proposed ELS dual-rail SRAM and previous dual-rail SRAMs.

	Interface	Array	Hybrid	ELS (proposed)
Low-power in SRAM	Not achieved	Achieved	Achieved	Achieved
Performance overhead (ns)	↓↓ (0.54)	↑↑ (1.61)	↑ (1.29)	↓ (0.600)
Read energy (fJ/cycle)	↑↑ (1.34)	↓ (1.08)	↓ (1.01)	↓ (1.05)
Write energy (fJ/cycle)	↑↑ (2.07)	↓ (1.35)	↓ (1.32)	↓ (1.37)
EDP (Normalized to SR)	↑ (1.56)	↑↑ (3.32)	↑↑ (2.54)	↓ (1.23)

SRAM, even though the energy consumption in the control block increases using VDDH.

A comparison of the total energy consumption of all SRAMs at the various VDDL levels is shown in Fig. 14. Although the total energy consumption of the proposed ELS dual-rail SRAM is slightly larger than that of the hybrid dual-rail SRAM, energy-saving compared to single-rail SRAM is notable as VDDL decreases. When VDDL is 0.422V, the proposed dual-rail SRAM shows 0.714-fold energy consumption compared to that of the single-rail SRAM, while the interface, array, and hybrid dual-rail SRAMs show 1.004-fold, 0.718-fold, and 0.678-fold energy consumption, respectively.

B. COMPARISON OF THE ENERGY SAVINGS AND EDP

The EDP comparison shown in Fig. 15 implies important merits of the proposed dual-rail SRAM. At all VDDL levels, the proposed ELS dual-rail SRAM achieves the smallest EDP among all dual-rail schemes. When VDDL is 0.422V, only 23% larger EDP than that of the single-rail SRAM is observed, while the EDP of the array and hybrid dual-rail SRAMs is 3.3 and 2.5 times that of the single-rail SRAM, respectively, due to the significant performance degradation. The EDP of the interface dual-rail SRAM is somewhat smaller than that of the two previous dual-rail SRAMs because the EDP is only affected by level shifters, neither achieving low-power nor suffering from significant performance degradation in the SRAM at the same time.

A comprehensive comparison between the proposed ELS dual-rail SRAM and the previous dual-rail SRAMs is summarized in Table 4.

C. AREA EVALUATION

The impact of the level-shifting sense amplifier and power-gated cross-coupled inverter based level-shifting write driver used in the proposed ELS dual-rail SRAM on area is evaluated. The evaluation is based on process information introduced in [27].

As the level-shifting sense amplifier includes the head switch and transmission gate access transistors which doesn't exist in the conventional VLSA [22]. The area of the level-shifting sense amplifier increases by 16.7% compared to that of the conventional VLSA. As the power-gated cross-coupled inverter based level-shifting write driver includes the level

TABLE 5. Processor configuration.

Common Parameter		
Processor	X86-based 1.5 GHz out-of-order, 128-entry reorder buffer, 64-entry load queue, 64-entry store queue	
Memory	DDR3-1600, 8 GB	
	Hybrid Dual-rail	Proposed Dual-rail
L1 I/D Cache	16 kB, 4-way set associative, 4-cycle tag/data/response latency	16 kB, 4-way set associative, 2-cycle tag/data/response latency
L2 cache	512 kB, 4-way set associative, 12-cycle tag/data/response latency	512 kB, 4-way set associative, 6-cycle tag/data/response latency

shifting circuit, the level shifters are not needed for input data in the ELS dual-rail SRAM. Furthermore, the cross-coupled PMOSs placed in every column in the hybrid dual-rail SRAM are not needed in the ELS dual-rail SRAM. As a result, the area of the write driver in the ELS dual-rail SRAM is smaller by 43% compared to that of the hybrid dual-rail SRAM.

Totally, the area of the sense amplifier and write driver in the ELS dual-rail SRAM is smaller by 32% than that of the hybrid dual-rail SRAM.

VI. PERFORMANCE EVALUATION OF ARCHITECTURE-LEVEL SIMULATION

To evaluate the architecture-level performance advantage of the proposed ELS dual-rail SRAM, an architecture-level simulation is performed using Gem5 simulator. Five INT and six FP benchmarks in SPEC2006 are selected to be simulated, as mentioned in Section II.

Table 5 lists the key processor configuration parameters used for the simulator. The hybrid dual-rail SRAM is selected for a comparison with the proposed ELS dual-rail SRAM owing to its relatively small performance overhead with low-power characteristic. Because the goal of the architectural simulation is to observe the impact of the performance of two dual-rail SRAMs on architecture-level performance, all parameters, except for the latencies of caches, are the same. Thus, only the performance difference of two dual-rail SRAMs affects the performance results of the simulation. The

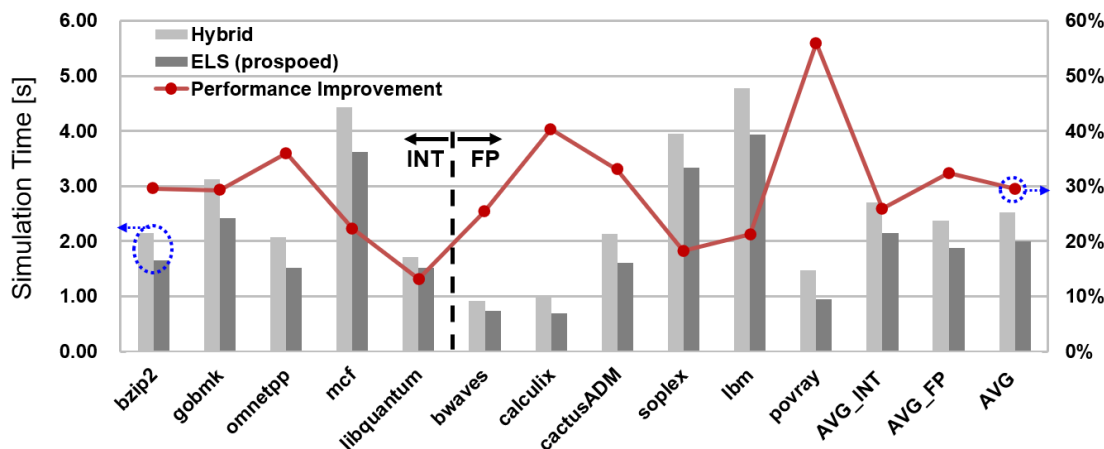


FIGURE 16. Performance evaluation results and performance improvement of proposed ELS dual-rail SRAM in SPEC2006 benchmarks.

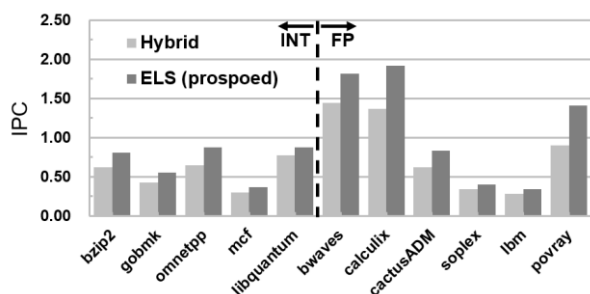


FIGURE 17. Instructions per cycle (IPC) simulation results.

L1 latencies of the proposed ELS dual-rail SRAM are set as 2-cycle because two CLK cycles are required to meet 0.6 ns read delay in Table 4, at a CLK frequency of 1.5 GHz. For the same reason, the L1 latencies of the hybrid dual-rail SRAM are set as 4-cycle to meet the 1.29 ns read delay.

The simulation results are shown in Fig. 16. For an accurate simulation, we simulate two billion instructions after fast-forwarding 16 billion instructions in all the selected benchmarks. The results of the five INT benchmarks are followed by those of the six FP benchmarks, and the average results of INT, FP, and all 11 benchmarks are shown at the end of Fig. 16, where the simulation time corresponds to performance, as the same number of instructions are used in simulation and fast-forwarding for each dual-rail SRAM. In addition, the instructions per cycle (IPC) results of each benchmark are summarized in Fig. 17. The polygonal line in Fig. 16 shows performance improvement in the proposed ELS dual-rail SRAM compared to the hybrid dual-rail SRAM. The embedded level-shifting dual-rail SRAM achieves shorter simulation time for all benchmarks, with an improvement of about 26% for INT benchmarks, 32% for FP benchmarks, and 29% on average.

VII. CONCLUSION

This paper proposes an embedded level-shifting (ELS) dual-rail SRAM to meet the low-power operation and high-performance demand in caches by improving the

energy-delay product with maximized energy-saving efficiency without significant performance degradation. In the proposed dual-rail SRAM, the proposed embedded level-shifting sense amplifier and output buffer dramatically reduce the performance degradation caused using a low supply voltage, while the energy consumption is significantly reduced using a low supply voltage to precharge BL. The read delay is also decreased by using a high supply voltage for the blocks and signals in the critical path, including the column MUX. Furthermore, the energy overhead caused by a large number of level shifters is reduced thanks to the proposed level-shifting write driver. In summary, with the circuit-level simulation results obtained for the 14-nm FinFET model, the proposed embedded level-shifting dual-rail SRAM shows 71.4% energy consumption with a performance overhead of 72% compared to the single-rail SRAM, while the hybrid dual-rail SRAM shows 67.8% energy consumption with a performance overhead of 270%. In architectural simulation with Gem5 simulator and SPEC2006 benchmarks, 29% performance improvement is achieved compared to the conventional hybrid dual-rail SRAM.

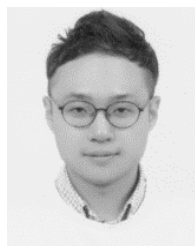
REFERENCES

- [1] R. G. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, and T. Mudge, "Near-threshold computing: Reclaiming Moore's law through energy efficient integrated circuits," *Proc. IEEE*, vol. 98, no. 2, pp. 253–266, Feb. 2010.
- [2] M. Powell, S.-H. Yang, B. Falsafi, K. Roy, and T. N. Vijaykumar, "Gated-vdd: A circuit technique to reduce leakage in deep-submicron cache memories," in *Proc. Int. Symp. Low power Electron. Design ISLPED*, 2000, pp. 90–95.
- [3] A. Sodani and C. Processor, "Race to exascale: Opportunities and challenges," in *Proc. Keynote Annu. IEEE/ACM 44th Annu. Int. Symp. Microarchitecture*, Dec. 2011, pp. 1–28.
- [4] J. Samandari-Rad and R. Hughey, "Power/energy minimization techniques for variability-aware high-performance 16-nm 6T-SRAM," *IEEE Access*, vol. 4, pp. 594–613, 2016.
- [5] M. Ling, X. Shang, S. Shen, T. Shao, and J. Yang, "Lowering the hit latencies of low voltage caches based on the cross-sensing timing speculation SRAM," *IEEE Access*, vol. 7, pp. 111649–111661, 2019.
- [6] P. Singh and S. Kumar Vishvakarma, "Ultra-low power high stability 8T SRAM for application in object tracking system," *IEEE Access*, vol. 6, pp. 2279–2290, 2018.

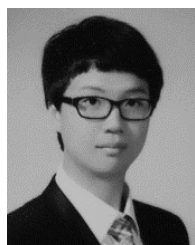
- [7] A. Ferrerón, D. Suarez-Gracia, J. Alastruey-Benede, T. Monreal-Arnal, and P. Ibanez, "Concertina: Squeezing in cache content to operate at near-threshold voltage," *IEEE Trans. Comput.*, vol. 65, no. 3, pp. 755–769, Mar. 2016.
- [8] Y. H. Chen, G. Chan, S. Y. Chou, H.-Y. Pan, J.-J. Wu, R. Lee, H. J. Liao, and H. Yamauchi, "A 0.6 v dual-rail compiler SRAM design on 45 nm CMOS technology with adaptive SRAM power for lower VDD_min VLSIs," *IEEE J. Solid-State Circuits*, vol. 44, no. 4, pp. 1209–1215, Apr. 2009.
- [9] M. Clinton, H. Cheng, H. Liao, R. Lee, C.-W. Wu, J. Yang, H.-T. Hsieh, F. Wu, J.-P. Yang, A. Katoch, A. Achyuthan, D. Mikan, B. Sheffield, and J. Chang, "12.3 a low-power and high-performance 10nm SRAM architecture for mobile applications," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2017, pp. 210–211.
- [10] J. Pille, C. Adams, T. Christensen, S. R. Cottier, S. Ehrenreich, F. Kono, D. Nelson, O. Takahashi, S. Tokito, O. Torreiter, O. Wagner, and D. Wendel, "Implementation of the cell broadband Engine in 65 nm SOI technology featuring dual power supply SRAM arrays supporting 6 GHz at 1.3 v," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 163–171, Jan. 2008.
- [11] J. Davis, D. Plass, P. Bunce, Y. Chan, A. Pelella, R. Joshi, A. Chen, W. Huott, T. Knips, P. Patel, K. Lo, and E. Fluhr, "A 5.6GHz 64kB dual-read data cache for the POWER6TM processor," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2006, pp. 2564–2571.
- [12] R. Joshi, Y. Chan, D. Plass, T. Charest, R. Freese, R. Sautter, W. Huott, U. Srinivasan, D. Rodko, P. Patel, P. Shephard, and T. Werner, "A low power and high performance SOI SRAM circuit design with improved cell stability," in *Proc. IEEE Int. SOI Conf.*, Oct. 2006, pp. 4–7.
- [13] M. Kim, I.-J. Chang, and H.-J. Lee, "Segmented tag cache: A novel cache organization for reducing dynamic read energy," *IEEE Trans. Comput.*, vol. 68, no. 10, pp. 1546–1552, Oct. 2019.
- [14] N. Binkert, B. Beckmann, G. Black, S. K. Reinhardt, A. Saidi, A. Basu, J. Hestness, D. R. Hower, T. Krishna, S. Sardashti, R. Sen, K. Sewell, M. Shoaib, N. Vaish, M. D. Hill, and D. A. Wood, "The gem5 simulator," *ACM SIGARCH Comput. Archit. News*, vol. 39, no. 2, pp. 1–7, May 2011.
- [15] J. L. Henning, "SPEC CPU2006 benchmark descriptions," *ACM SIGARCH Comput. Archit. News*, vol. 34, no. 4, pp. 1–17, Sep. 2006.
- [16] A. Phansalkar, A. Joshi, and L. K. John, "Analysis of redundancy and application balance in the SPEC CPU2006 benchmark suite," in *Proc. 34th Annu. Int. Symp. Comput. Archit. ISCA*, 2007, pp. 412–423.
- [17] T. Song, W. Rim, S. Park, Y. Kim, G. Yang, H. Kim, S. Baek, J. Jung, B. Kwon, S. Cho, H. Jung, Y. Choo, and J. Choi, "A 10 nm FinFET 128 mb SRAM with assist adjustment system for power, performance, and area optimization," *IEEE J. Solid-State Circuits*, vol. 52, no. 1, pp. 240–249, Jan. 2017.
- [18] Y. Ho, S.-Y. Hsu, and C.-Y. Lee, "A variation-tolerant subthreshold to superthreshold level shifter for heterogeneous interfaces," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 2, pp. 161–165, Feb. 2016.
- [19] J. Wang, S. Nalam, and B. H. Kalhoun, "Analyzing static and dynamic write margin for nanometer SRAMs," in *Proc. ACM/IEEE Int. Symp. Low Power Electron. Design (ISLPED)*, Aug. 2008, pp. 129–134.
- [20] M. H. Abu-Rahma and M. Anis, *Nanometer Variation-Tolerant SRAM: Circuits and Statistical Design for Yield*. Springer, 2013.
- [21] T. Song et al., "12.2 a 7nm FinFET SRAM macro using EUV lithography for peripheral repair analysis," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2017, pp. 208–209.
- [22] T. Na, S.-H. Woo, J. Kim, H. Jeong, and S.-O. Jung, "Comparative study of various latch-type sense amplifiers," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 2, pp. 425–429, Feb. 2014.
- [23] C.-H. Jan et al., "A 14 nm SoC platform technology featuring 2nd generation tri-gate transistors, 70 nm gate pitch, 52 nm metal pitch, and 0.0499 um² SRAM cells, optimized for low power, high performance and high density SoC products," in *Proc. Symp. VLSI Technol. (VLSI Technol.)*, Jun. 2015, pp. T12–T13.
- [24] S. Natarajan et al., "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 um² SRAM cell size," in *IEDM Tech. Dig.*, Dec. 2014, pp. 3.7.1–3.7.3.
- [25] H. Kawasaki et al., "Challenges and solutions of FinFET integration in an SRAM cell and a logic circuit for 22 nm node and beyond," in *IEDM Tech. Dig.*, Dec. 2009, pp. 1–4.
- [26] M. D. Giles, N. A. Radhakrishna, D. Becher, A. Kornfeld, K. Maurice, S. Mudanai, S. Natarajan, P. Newman, P. Packan, and T. Rakshit, "High sigma measurement of random threshold voltage variation in 14nm logic FinFET technology," in *Proc. Symp. VLSI Technol. (VLSI Technol.)*, Jun. 2015, pp. T150–T151.
- [27] K. Fischer et al., "Low-k interconnect stack with multi-layer air gap and tri-metal-insulator-metal capacitors for 14nm high volume manufacturing," in *Proc. IEEE Int. Interconnect Technol. Conf. IEEE Mater. Adv. Metallization Conf. (IITC/MAM)*, May 2015, pp. 5–8.
- [28] T. Na, J. Kim, J. P. Kim, S. H. Kang, and S.-O. Jung, "An offset-canceling triple-stage sensing circuit for deep submicrometer STT-RAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 7, pp. 1620–1624, Jul. 2014.
- [29] S.-C. Luo, C.-J. Huang, and Y.-H. Chu, "A wide-range level shifter using a modified wilson current mirror hybrid buffer," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 6, pp. 1656–1665, Jun. 2014.
- [30] S. Lutkemeier and U. Ruckert, "A subthreshold to above-threshold level shifter comprising a wilson current mirror," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 9, pp. 721–724, Sep. 2010.
- [31] R. Matsuzuka, T. Hirose, Y. Shizuku, N. Kuroki, and M. Numa, "A 0.19-V minimum input low energy level shifter for extremely low-voltage VLSIs," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 2948–2951.
- [32] H. Jeong, Y. Yang, J. Lee, J. Kim, and S.-O. Jung, "One-sided static noise margin and Gaussian-Tail-Fitting method for SRAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 6, pp. 1262–1269, Jun. 2014.



TAE HYUN KIM was born in Busan, South Korea, in 1991. He received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2016, where he is currently pursuing the Ph.D. degree in electrical and electronic engineering. His current research interests include low-power FinFET-based static random access memory (SRAM) design, parasitic component aware SRAM peripheral circuit design, and high reliability cross-point array memory design.



HANWOOL JEONG was born in Seoul, South Korea, in 1987. He received the B.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, in 2012 and 2017, respectively. From 2017 to 2019, he was with the Foundry Division, Samsung Electronics Company Ltd., Hwaseong, South Korea, where he was involved with the circuit design and verification of 4nm/5nm memory compiler. Since 2019, he has been a Professor with Kwangwoon University, Seoul. His current research interests include memory circuit design, low-voltage/low power digital logic and neuromorphic, and machine learning circuit design.



JUHYUN PARK (Member, IEEE) was born in Incheon, South Korea, in 1988. He received the B.S. degree in electronic and electrical engineering from Hongik University, Seoul, South Korea, in 2012, and the Ph.D. degree in electrical and electronic engineering from Yonsei University, Seoul, in 2020. He joined SK Hynix, Icheon, in 2020, where he is involved with mobile DRAM design.



power circuit design for mobile applications and embedded SRAM memory.

HOONKI KIM (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2007 and 2013, respectively. From 2013 to 2015, he was a Postdoctoral Associate with the University of Minnesota, Minneapolis, MN, USA. He joined Samsung Electronics, Hwaseong, South Korea, in 2015, where he is involved with design technology co-optimization (DTCO) and SRAM design. His research interests include high-speed low-



TAEJOONG SONG (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, in 2010. He joined Samsung Electronics, in 1997, where he is leading the Library IP Group of Standard Cell, Memory Compiler, IO, OTP in System LSI, Samsung Electronics. His research interest includes design technology co-optimization (DTCO) for high-speed/low-power in 10/7 nm FinFET technology.



graphic RAM, and window RAM, and at T-RAM Inc., Mountain View, CA, USA, where he was the Leader of the Thyristor-Based Memory Circuit Design Team. From 2003 to 2006, he worked with Qualcomm Inc., San Diego, CA, USA, where he was involved with high-performance low-power embedded memories, process variation tolerant circuit design, and low-power circuit techniques. Since 2006, he has been a Professor with Yonsei University. His current research interests include process variation-tolerant circuit design, low-power circuit design, mixed mode circuit design, and future generation memory and technology.

SEONG-OOK JUNG (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Yonsei University, Seoul, South Korea, in 1987 and 1989, respectively, and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2002. From 1989 to 1998, he worked with Samsung Electronics Company Ltd., Hwasung, South Korea, where he was involved with specialty memories, such as video RAM,

• • •