Robust Ultra-Low Voltage ROM Design

Mingoo Seok, Scott Hanson, Jae-Sun Seo, Dennis Sylvester, David Blaauw

Electrical Engineering and Computer Science
University of Michigan, Ann Arbor
{mgseok@ hansons@ jseo@ dmcs@ blaauw@} umich.edu

Abstract— SRAM dominates standby power consumption in many systems since the power supply cannot be gated as in logic blocks. The use of ROM for parts of instruction memory can alleviate this power bottleneck in mobile sensing applications such as implantable biomedical and environmental sensing systems, which can spend up to 99% of their lifetimes in standby mode. However, robust ROM design becomes challenging as the supply voltage is reduced aggressively. In this paper, three different ROM topologies are investigated and compared for ultra-low voltage operation. A simple method to estimate the theoretical robustness at low voltage is proposed and applied to the ROM topologies. A test circuit fabricated in a carefully-selected 0.18μm CMOS technology reveals that our proposed static NAND ROM structure improves performance by 26X, energy by 3.8X and lowest functional supply voltage by 100mV over a conventional dynamic NAND ROM.

I. INTRODUCTION

Applications that depend on energy scavenging or on-die batteries for a power source require ultra–low power consumption to guarantee long lifetime. Such applications range from implantable biomedical systems to environmental sensing systems among others. Scaling supply voltage near or below the device threshold voltage (V_{th}) has recently emerged as an attractive method to save switching energy [1][2]. In addition to minimizing switching energy, standby energy reduction is particularly critical since sensing systems can spend >99% of their lifetimes in standby mode [3].
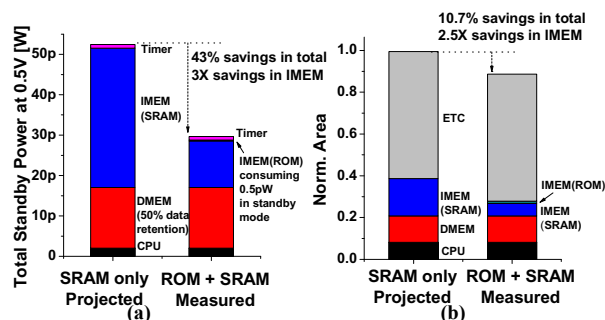

Figure 1. (a) Power and (b) area comparisons for SRAM-only IMEM (projected) and an SRAM/ROM hybrid IMEM (measured).

The authors of [4] observe that SRAM (Static Random Access Memory) for both data and instructions is the dominant source of total standby energy consumption in a typical sensing platform. It is therefore paramount to minimize the standby power consumption of the memories. Although most data memory (DMEM) must be both read and written, instruction memory (IMEM) can be re-optimized to take advantage of its read-only nature. For example, by storing common procedures in ROM (Read Only Memory) with a power gating switch designed for ultra-low voltage operation, both standby power and area can be reduced. Figure 1 shows that standby power can be reduced by 43% and area reduced by 10.7% in a sensing platform by replacing 128 out of 192 SRAM words with power-gated ROM.

However, there are four key challenges for designing robust ROM at ultra-low voltages: 1) The reduced on-current to off-current ratio causes robustness problems, 2) there is potentially a large skew in

beta ratio (relative strength between NFET and PFET) at low voltage, 3) for dynamic ROM styles, conventional keepers (half-latches) are likely to lose state and 4) significant variability further complicates each of the previous three issues.

In this paper, we explore the design of ultra-low voltage ROM. First we investigate the challenges of designing conventional dynamic NAND ROM at ultra-low voltages and propose circuit techniques to overcome these challenges. We also propose a back-of-the-envelope method, referred to as a *current margin plot*, which estimates the theoretical functionality of ROM at ultra-low voltages and provides guidelines for design decisions. We then propose two alternative ROM topologies, static NAND ROM and static NAND-NOR ROM, that improve robustness, performance, and energy-per-operation compared to dynamic NAND ROM. The current margin plot is used to estimate robustness for the two static ROM topologies. We conclude by describing a 0.18μm test chip that includes structures for each of the three ROM topologies discussed. The 0.18μm technology is chosen due to a superior balance between switching and leakage energy relative to more recent technologies. Measurements show that the static NAND ROM improves performance by 26X, energy by 3.8X, and minimum functional supply voltage by 100mV over a conventional dynamic NAND ROM.

II. DYNAMIC NAND ROM DESIGN

This section first investigates the challenges of ultra-low voltage ROM design with a particular focus on the dynamic NAND ROM topology (Figure 2(a)), which is commonly used in superthreshold operation. Although dynamic NOR is also commonly used in superthreshold regime, it is not considered in this study due to the reason discussed in the section II.C. Then a method called a current margin plot is proposed to show theoretical robustness at ultra-low voltages, which can be applied to any ROM topology. Using this method we describe the design of a dynamic NAND ROM targeting ultra-low voltage operation.

A. Challenges of dynamic NAND ROM design at low voltages

The dynamic NAND ROM (Figure 3(a)) operates in two phases: precharge and evaluation. In precharge when clock is low, the dynamic node is charged up to V_{dd}. In evaluation when the clock is high, the dynamic node is either discharged to V_{ss} by stacked NFETs or held at V_{dd} by a half-latch keeper depending on the read word line signals. Having an NFET for a specific read word line means a high output value since the NFET for the word line is turned off.

Operation becomes less robust in the ultra-low voltage regime for several reasons. First, the on- to off-current ratio is reduced (Figure 3(a)), resulting in robustness problems since on-current becomes less distinguishable from off-current. This problem is exacerbated in ROM design since ROM usually has a large number of FETs in series for NAND, or in parallel for NOR styles [5]. The FETs in series limit on-current while FETs in parallel increase the worst-case leakage current. As shown in Figure 3(b), the on-current decreases super-linearly as more FETs are connected in series, resulting in a worse on- to off-current ratio. In the technology used in this work, the on- to off-current ratio of 32-stacked NFETs is only ~152X at 0.5V, which is several orders of magnitude smaller than at nominal voltage.
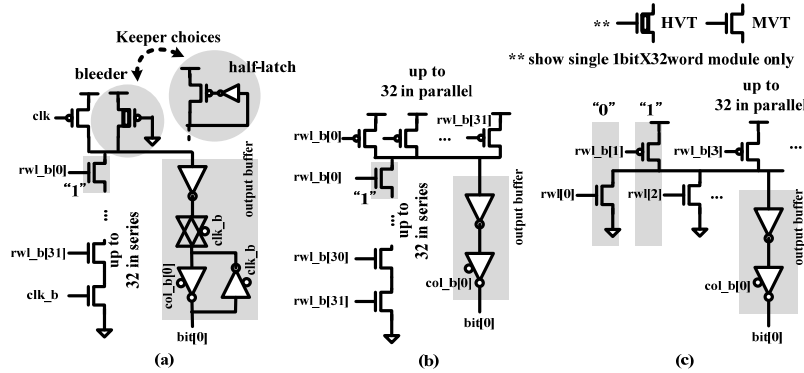
Figure 2. Schematics of three ROMs for ultra-low voltage: (a) dynamic NAND, (b) static NAND, (c) static NAND-NOR

Additionally, the large skew in beta ratio can further aggravate this on- to off-current ratio problem. Beta ratio can be dramatically different between subthreshold and superthreshold since on-current is exponentially dependent on V_{th} and devices are typically optimized for superthreshold operation. In this technology, the min-sized NFET is ~20X stronger than the min-sized PFET at V_{dd}=0.5V, compared to 2.7X at V_{dd}=1.8V, as shown in Figure 3(a). Therefore the ratio of PFET on-current to NFET off-current is reduced by roughly 20X in addition to the already-reduced on-current to off-current ratio of ultra-low voltage operation.
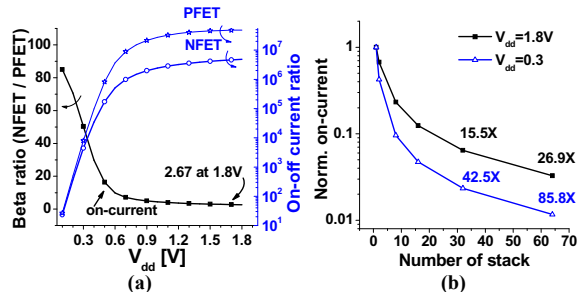

Figure 3. (a) Beta ratio and on- to off-current ratio, (b) on-current reduction over number of stack (for minimum-sized FET)

The functionality of the half-latch keeper (Figure 2(a)) in the ultra-low voltage regime is another problem for dynamic ROM design. The half-latch becomes more important at low voltages since the charge on dynamic nodes is reduced linearly with scaled supply voltage while leakage current stays almost constant. The half-latch is not able to maintain its state for two reasons. First, its retention ability is reduced. When we view the half-latch keeper as broken back-to-back inverters in SRAM, its static noise margin is known to degrade in ultra-low voltage regime. Second, the same amount of charge sharing has a more destructive effect at low voltages. Finally, large variability further complicates low voltage dynamic NAND ROM design. Simulations show that if NFET off-current is larger than nominal by 1σ and PFET on-current smaller than nominal by 1σ due to process variations, the total on- to off-current ratio is reduced by 4.8X at 0.5V.

B. On-current to off-current plot

In this section a back-of-the envelope method, a "current margin plot", is described to estimate the theoretical robustness of ROM in the ultra-low voltage regime. All the factors mentioned in the previous section are accounted for in the method. Here we apply it to a 32-stack dynamic NAND ROM with a high-V_{th} PFET bleeder with length of 0.45µm and width of 0.33µm (Figure 2(a)).

Figure 4 shows the margin of on- to off-current ratio in the 32-stack dynamic NAND ROM with bleeder for two different operations at different voltages: 1) evaluation of a one (eval-1) and 2) evaluation of a zero (eval-0). The eval-1 (left side) is the case where the output is maintained by the bleeder. Here the worst case off-current through the NFET stack should be smaller than the current that the bleeder provides to guarantee functionality. A guardband equal to the standard deviation of off-current is included to estimate worst-case off-current and is denoted as VAR in Figure 4. On the other hand, complete discharge through stacked NFETs is required for eval-0 (right side of Figure 4). Here the discharging current through NFETs should be larger than the bleeder current. However the discharging current is reduced by the series connection (STK) and guardband for process variation (VAR), resulting in a range of just 20X between the stack and the bleeder at 0.5V.
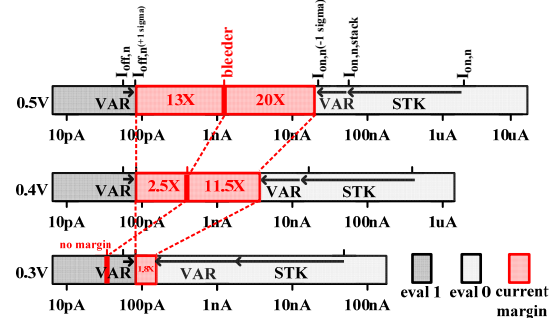

Figure 4. Current margin plot for 32-stack dynamic NAND ROM with HVT bleeder

The magnitude of the guardbands for FET current due to process variation (VAR) is determined based on the standard deviation taken from 1000 Monte Carlo simulations. Here, the log of current is assumed as a Gaussian distribution, which is based on the fact that subthreshold current is an exponential function of normally distributed V_{th}, which is the dominant factor for current variation [6] at ultra-low voltage. One thing to note here is that variation in off-current is almost constant while on-current variation is increased with smaller supply voltage, which is considered in the current margin plot. We set VAR based on the current at μ±1σ of V_{th}.[1] The current reduction due to the large stack (STK) is based on simulation results (Figure 3(b)).

The margin of on- to off-current ratio provides information about two circuit metrics: robustness and performance. Clearly larger margins offer more robustness in light of substantial process variation. In addition to robustness, the margin dictates circuit performance. For instance, a large margin for the eval-1 case implies fast recovery from signal degradation of dynamic nodes. A large margin between dis-

---

[1] A single standard deviation is used here due to the assumption of small designs at ultra-low voltages (e.g., sensor processors) but a more conservative value could also be employed.

16-6-2

424

charging current and bleeder current is preferred for reduced contention.

The margin of the on- to off-current ratio is diminished as the supply voltage is scaled down. Since process variation in the bleeder current can further degrade the margin, robustness at ultra-low voltage for this ROM topology is questionable.

C. A 32-stack dynamic NAND ROM with HVT bleeder

In the previous section, a 32-stack dynamic NAND ROM with a bleeder is used to illustrate the current margin plot. This section discusses the design decisions regarding ROM topology, stack height and keeper style, which collectively point to a 32-stack dynamic NAND ROM with bleeder as a reasonable design choice.
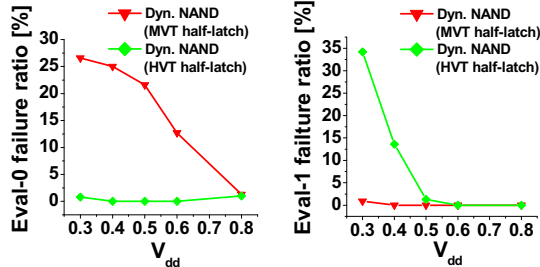

Figure 5. Failure rate for the dynamic NAND with half-latches. (1000 Monte Carlo iterations with die-to-die and mismatch variations)

First, dynamic NAND is chosen over dynamic NOR ROM due to the large skew in beta ratio. Consider a 32-stack dynamic NOR ROM. The off-current through 31 parallel-connected NFETs is only ~20X smaller than the on-current of a single PFET serving as keeper at V_{dd}=0.5V, even without considering process variation. In addition, this reduced on- to off-current ratio degrades performance, which is one of the primary advantages of the NOR topology. Larger leakage power and larger footprint compared to the NAND topology are other drawbacks. Therefore the dynamic NAND structure is chosen over a NOR topology in this study. Note that this decision is motivated primarily by technology limitations. A different technology may make the NOR structure more attractive.

Second, given a dynamic NAND structure, we select a stack height of 32. A taller stack reduces area but can cause robustness problems due to small discharging currents and significant charge sharing. A stack of 64 devices would reduce the margin of on- to off-current ratio between the bleeder current and the worst case on-current by 2X compared to a stack of 32 devices.

Finally, a HVT bleeder is chosen over the half-latch keeper configuration. Monte Carlo simulation with die-to-die and mismatch variations shows that two half-latches with different strengths, (a medium V_{th} (MVT) device with W/L=0.33 μm/1.8μm and an HVT device with W/L= 0.33μm/0.45μm) fail to discharge (eval-0) or hold (eval-1) dynamic nodes in the ultra-low voltage regime as shown in Figure 5. The MVT half-latch often becomes so strong that series-connected NFETs are unable to discharge the dynamic node while the HVT half-latch is often unable to supply enough charge to overcome charge sharing. However the HVT bleeder operates more robustly than the half-latch in the ultra-low voltage regime. An HVT bleeder provides the same on-current as the HVT half-latch for eval-0, so the series connected NFETs are able to discharge the dynamic node. Additionally, the bleeder constantly provides current even after the dynamic node accidentally pulls low, so the correct value will eventually be restored in contrast to a half-latch.

Setting the appropriate strength of the bleeder is important due to the tradeoff between recovery and contention. The margin of on- to off-current in Figure 4 specifies the available strength that the bleeder can have. If the bleeder strength resides outside the margin, it can cause incomplete discharge for eval-0 or poor recovery for eval-1. However setting bleeder strength is a non-trivial task. While keeper strength is adjusted at nominal V_{dd} through gate sizing, it cannot be applied in an area efficient manner in the ultra-low voltage regime due to the exponential variability in current. As shown in Figure 4, if a MVT device is used as a bleeder, the on-current is reduced by ~3 orders of magnitude to reside in the allowed margin, requiring infeasible length biasing. Therefore other knobs such as using different V_{th} devices or applying body bias should be considered. We use a different V_{th} in this work to avoid generating an extra body bias.

III. STATIC NAND AND NAND-NOR ROM

Although we have shown that dynamic NAND ROM can operate at very low voltages, the performance, energy-per-operation and minimum functional voltage are unsatisfactory due to the small current margin. The bleeder is among the most important components of this design style, which gives rise to a challenging sizing tradeoff. Therefore, new topologies without a bleeder are worth investigating. In this section, we describe the design of 2 full-static ROMs: 32-stack NAND and 32-leg NAND-NOR as shown in Figures 2(b) and 2(c). Since these structures have no bleeder, a larger on- to off-current margin is expected.

A. Investigating static ROM topologies

This section applies the current margin plot to the two static ROM topologies to investigate theoretical robustness at ultra-low voltages. Figure 6 shows the margin of on- to off-current ratio for the 32-stack static NAND ROM. Only the eval-0 case is considered here since it is the most stringent for large NFET stacks. Larger margin is still observed after incorporating the effect of parallel connection of PFETs (denoted as PAR), series connection of NFETs, and guardbands for process variation. The avoidance of the bleeder also helps increase the margin and ease design. Overall a 17X margin is maintained at a very aggressive V_{dd} of 0.3V, compared to zero margin for the 32-stack dynamic NAND ROM described earlier.
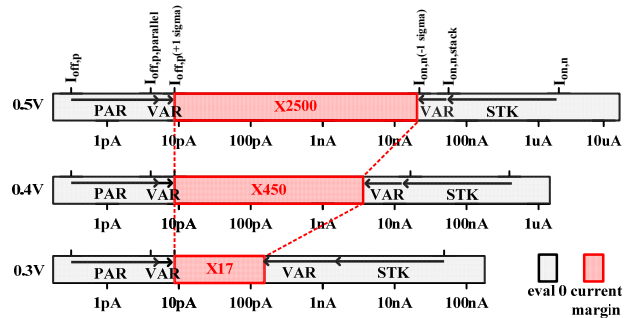

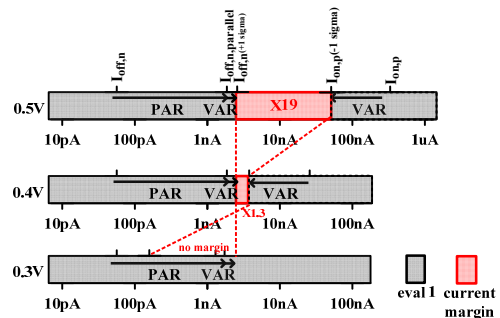Figure 6. Current margin plot for 32-stack static NAND ROM

Figure 7. Current margin plot for 32-leg static NAND-NOR ROM

If NFET and PFET strengths are balanced, the static NAND ROM can be improved by replacing the long stack with parallel legs using

16-6-3

425

inverted input signals as shown in Figure 2(c). However since the technology used in this study has a large beta ratio at low voltages, this topology is less robust than the static NAND ROM. Figure 7 shows the much reduced on-off margin in the eval-1 case where a single PFET contends with 31 NFETs, which is same as in a dynamic NOR topology. The current margin disappears at 0.3V, as in the dynamic NAND ROM. However, better robustness is expected over dynamic NAND ROM since no bleeder is present. Although the NAND-NOR ROM topology is not ideal in the technology used in this study, it may be a good candidate for technologies with more balanced beta ratios at low voltage.

B. Static NAND ROM Monte Carlo Analysis

Since the current margin plot is a first-order method to estimate robustness, Monte Carlo simulations considering all sources of variation are performed to investigate the effectiveness of the plot as well as the robustness of the ROM topologies. As shown in Figure 8, the static NAND ROM starts to fail at 0.3V in the eval-0 case due to the large NFET stack. The failure voltage is higher than that estimated by the current margin plot since the latter considers only one standard deviation of variation. In comparing topologies, the minimum functional voltage for the static NAND ROM is larger than that for the dynamic NAND ROM by nearly 200mV, confirming that the current margin plot is able to track the trends as well as the static NAND ROM is more robust.
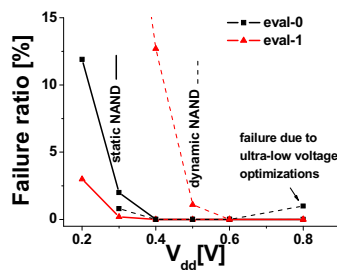

Figure 8. 1000 Monte Carlo SPICE simulations for two ROM topologies considering mismatch and die-to-die

IV. Measurement Results

A 10x128bit dynamic NAND ROM with HVT bleeder, a 10x128bit static NAND ROM, and a 10x128bit static NAND-NOR ROM were fabricated in a 0.18μm CMOS technology. Each ROM contains an identical set of random data patterns as well as patterns causing worst case charge sharing. The worst case pattern for the dynamic NAND ROM and the static NAND ROM is 31 series-connected NFETs while the worst case for the static NAND-NOR structure is a single PFET connected to 31 parallel-connected NFETs. Relevant silicon measurements and dimensions are shown in Figure 10.

The two static topologies show dramatically improved maximum operating frequency, energy-per-operation, and minimum functional voltage compared to the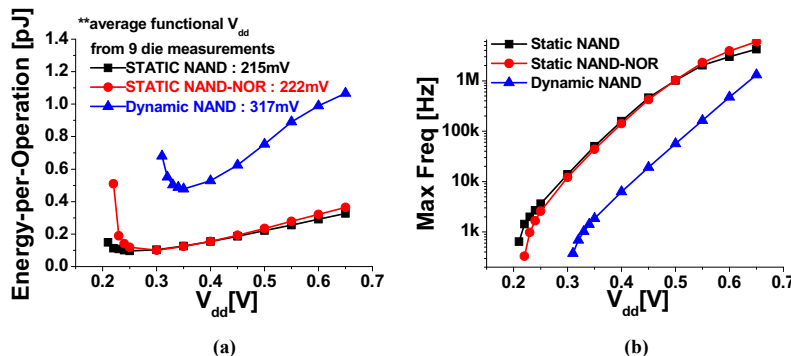 dynamic NAND ROM. The small on- to off-current ratio degrades performance in the dynamic NAND ROM, leading to substantially lower performance. The static NAND ROM and the static NAND-NOR ROM show similar energy, performance and minimum operating voltage numbers, though the static NAND ROM has a small advantage over the static NAND-NOR ROM, as predicted in previous sections.

Figure 9 shows the effect of variability on the performance of the static NAND ROM. The variation in maximum operating frequency (σ/μ) across 20 dies at 0.35V is just 18%. This number falls between the bounds set by a Monte Carlo simulation that includes die-to-die and mismatch variation and another Monte Carlo simulation considering mismatch only. Since all the 20 dies come from a single wafer, it is not surprising that the relatively small measured variability is much closer to the variability predicted by the mismatch-only simulation.
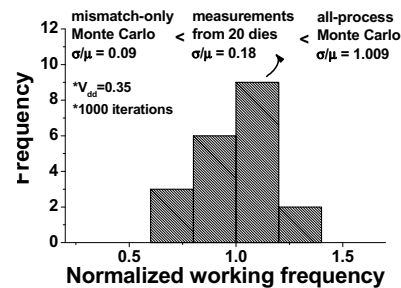

Figure 9. Histogram of operating frequency of static NAND ROM

V. Conclusions

Three different ROM topologies for ultra-low voltage operation are investigated with the test chip fabrication in an industrial 0.18μm CMOS technology. The challenges in ultra-low voltage design are analyzed and incorporated in the current margin plot which is devised for estimating theoretical low voltage robustness. Silicon measurements shows that the static NAND ROM shows 26X faster performance, 3.8X smaller energy-per-operation and 100mV smaller minimum working voltage than the dynamic NAND ROM with 33% area penalty. The static NAND-NOR ROM is also studied as a potential candidate for other technologies.

References

[1] A. Wang, *et al*, "A 180mV FFT Processor Using Subthreshold Circuit Techniques", ISSCC, 2004

[2] M. Hwang, *et al*, "A 85mV 40nW Process-Tolerant Subthreshold 8x8 FIR Filter in 130nm Technology", Symposium on VLSI Circuits, 2007

[3] L. Nazhadili, *et al*, "SenseBench: toward on accurate evaluation of sensor network processors", Workload Characterization Symposium, 2005

[4] M. Seok, *et al*, "The Phoenix Processor: A 30pW Platform for Sensor Applications", Symposium on VLSI Circuits, to be published in 2008

[5] S. Hsu, *et al*, "A 9GHz 320x80bit Low Leakage Microcode Read Only Memories in 65nm CMOS", ESSCIRC, 2006

[6] N. Verma, et al, "Nanometer MOSFET Variation in Minimum Energy Subthreshold Circuits", TED, vol. 55, no. 1, Jan, 2008

Figure 10. (a) Measured energy-per-operation (b) measured frequency measurement (c) die photo and dimensions