WHITE PAPER

Write Assist in Low-Voltage SRAMs

Vikas Chandra
Principal Design Engineer

Pratul Sharma
Technical Marketing Manager

May 2012

Abstract

Write assist techniques are now commonly used to lower the minimum operating voltage (Vmin) of an SRAM. One of the key reasons to push the SRAM Vmin lower is to enable efficient Dynamic Voltage and Frequency Scaling (DVFS) to save power. In this paper, we discuss the basics of SRAM faliure mechanisms, fundamentals of write assist techniques, ARM® Artisan® Low Voltage Memory Compilers with the write assist feature and results from GLOBALFOUNDRIES 28nm-SLP memories.

Contents

Introduction

Static random access memory (SRAM) is a critical part of most VLSI system-on-chip (SoC) applications. The SRAM bit cell design has to cope with stringent requirement on the cell area leading to minimum (or close to minimum) sized transistors. This increases the impact of local mismatch between the bitcell transistors and leads to functional issues like reduced stability and write-ability, especially at lower supply voltages. Due to the process scaling trend, device variations and leakage are increasing sharply with each shrinking technology node. As the need for low power systems grows, the supply voltage (VDD) has been scaled down to reduce both dynamic and leakage power. The operation of the SRAM at lower supply voltage becomes even more challenging. The predominant yield loss from increased device variability occurs at minimum operating voltage, a term defined as Vmin. The failures at Vmin can be due to write failure, read disturb failure, access failure or retention failure. Further, as SRAM capacity continues to increase, variability degrades both stability and write-ability in the array. Thus, scaling SRAM voltage to design a low power system becomes challenging due to the trade-off with cell stability and write-ability.

One of the key reasons to push the SRAM Vmin lower is to enable DVFS (Dynamic Voltage and Frequency Scaling) to save power. DVFS is a power management scheme which jointly optimizes performance and power consumption for energy constrained applications. The main idea is to reduce the supply voltage (and operating frequency) when the design is not doing critical tasks. This leads to significant savings in power consumption (both dynamic and leakage). Hence a DVFS scheme would like to push the voltage and frequency as low as possible. However, typically, the lower limit of Vmin is set by SRAM arrays and hence the supply voltage of the whole system cannot go lower than that. Hence it is imperative to push the SRAM Vmin lower so that the DVFS scheme can be made more efficient.

Failure Mechanisms

Usually the SRAM Vmin is limited by either write failure or read disturb failure. It is, however, difficult to predict a priori as to which of the two failure mode dominates because it is dependent upon many factors including the bitcell architecture, technology node etc. In this paper we will focus on write failure only. Write failure is defined as the failure to intentionally flip the value of the bit cell during the write operation. Various write assist (WA) schemes have been proposed in literature to help the bit cell to flip during the write. For the SRAMs in which the Vmin is dictated by write failures, the WA techniques push the Vmin lower and make them more amenable to voltage scaling.

Write Assist

With technology scaling, it is becoming difficult to write to SRAMs even at nominal supply voltage and the challenges become more apparent at lower supply voltages. In this paper, we will quantify SRAM write-ability using a metric called critical WL pulse width (or WL_{crit}). WL_{crit} is defined as the minimum wordline pulse width which is needed to flip the bit cell during write operation. The lower the value of WL_{crit}, the easier it is to write the cell and vice-versa. Figure 1 shows the trend of WL_{crit} with respect to supply voltage.

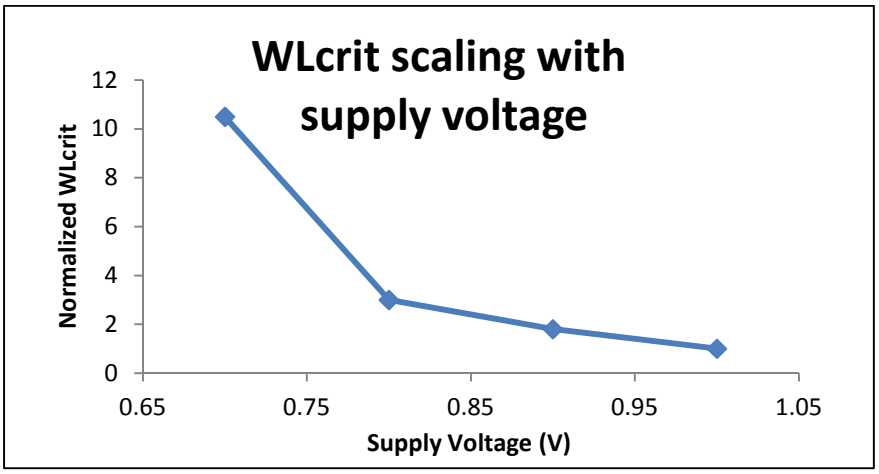

Figure 1: Change in WL_{crit} with voltage scaling

The increase in WL_{crit} is ~10X as the voltage scales down from 1V to 0.7V. This trend is very alarming since supply voltage is frequently dynamically scaled in System-on-Chip designs to reduce power consumption. The substantial increase in WL_{crit} will increase the Vmin of the SRAM and limit its applicability in low power designs. Hence, there is a need to improve the write performance at low supply voltages. The techniques which aid the bit cell in changing the state during write operation are called write assist (WA) techniques ([5], [7], [8], [9], [10]) and now they are widely used in most low power SRAMs.

In essence, WA techniques aid the bit cell to flip state during the write operation. There are many ways to implement write-assist - few notables ones are wordline boosting, negative bitline,

VDD lowering and VSS raising. Each technique has its own merits and demerits. In this paper we focus on the wordline boosting scheme which assists the bit cell to flip the state during a write event by boosting the word-line higher than the supply voltage (Figure 2). The boosting increases the Vgs of the access transistor and hence increases its drive strength. The increased drive strength of the access transistor aids significantly in flipping the bit cell.

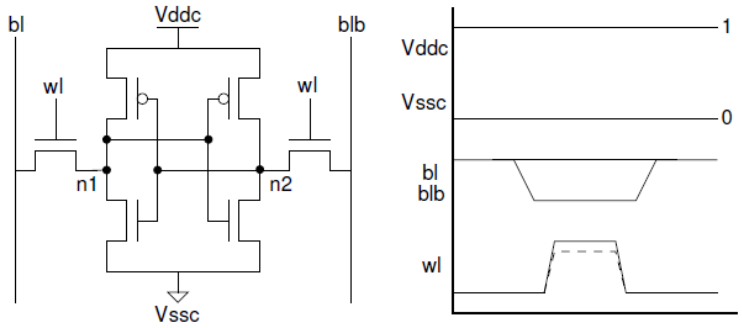

Figure 2: Write assist based on word-line boosting

Figure 3 shows the impact of word-line boosting based WA on the WLcrit. The WLcrit in this case is substantially better than the nominal case with no WA. As shown in Figure 3, the benefits of this WA scheme increases significantly as the supply voltage is scaled down, thus decreasing the SRAM Vmin.

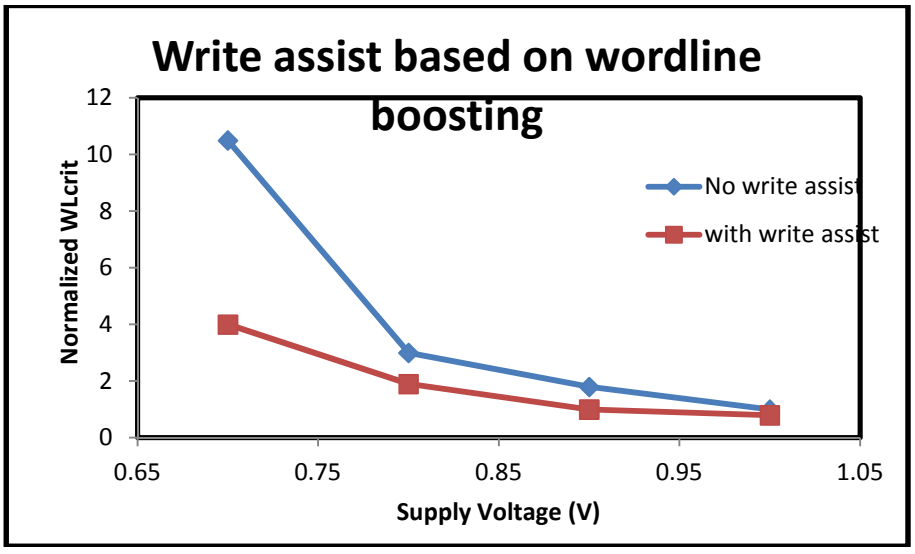

Figure 3: Normalized WL_{crit} for write assist based on word-line boosting

Write Assist and Impact of Process Variations

The significance and complexity of process variations is increasing with technology scaling. In sub-40nm technology, these variations can be classified into two groups, based on the

mechanism of the variation - systematic and random. These sources of variations severely limit the ability to push the performance of a design. Figure 4 shows the distribution of WL_{crit} for a bit cell with no write assist and Figure 5 shows the distribution of WL_{crit} for a bit cell with wordline boosting write assist.

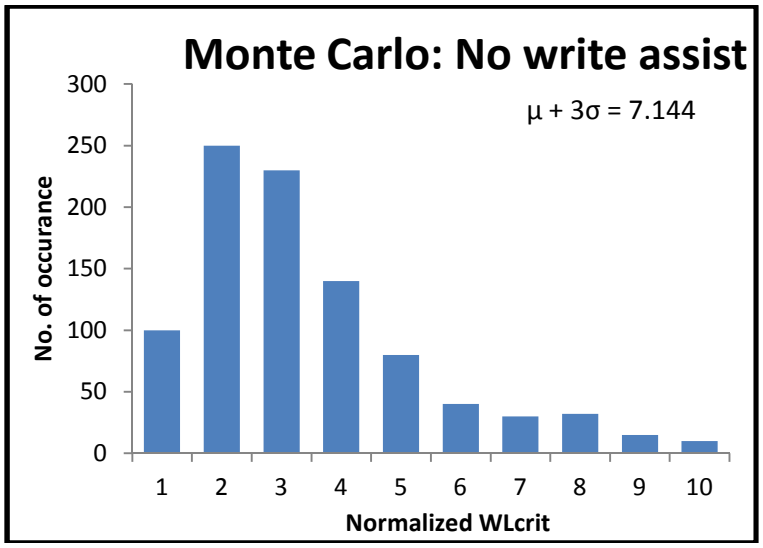

Figure 4: Statistical spread of WLcrit (no write assist)

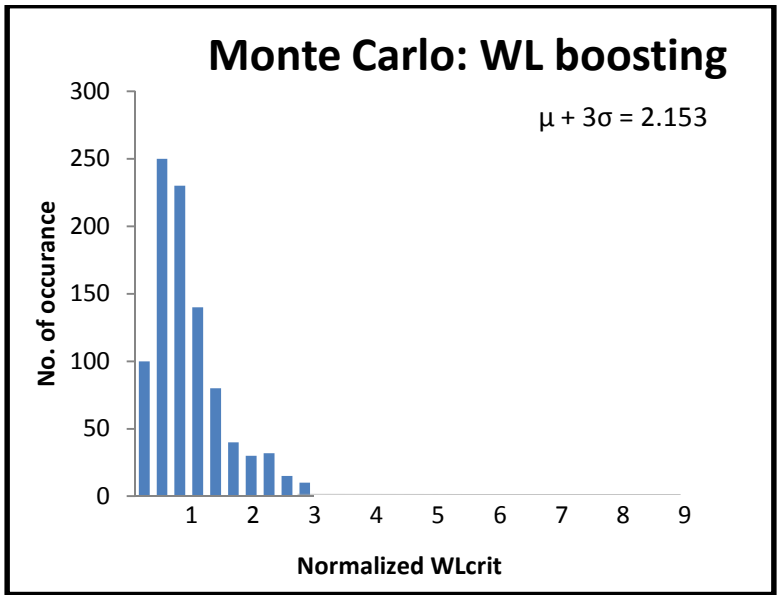

Figure 5: Statistical spread of WLcrit (WL boosting WA)

It can be noted from Figures 4 and 5 that write assist not only cuts down on the mean WL_{crit} but also the spread of the distribution. This essentially means that the margins required for the write

assist based implementation will be significantly less and the hence the design will have more performance as compared with the non write assist implementation.

Benefits of Write Assist

As discussed earlier write assist allows memories to run at a lower voltage. A benchmark on GLOBALFOUNDRIES 28nm (28SLP) low-power memories (at worst case corner) shows:

- Savings of up to 14% in dynamic power for the iso-frequency operation
- Savings of up to 25% in leakage power
- Savings of up to 35% in dynamic power with DVFS

These savings are achieved when write assist is enabled and the memories are run and lower than standard nominal voltage.

Figure 6 shows the dynamic power savings and Figure 7 shows the leakage power savings using write assist for 7 different sizes of memory instances.

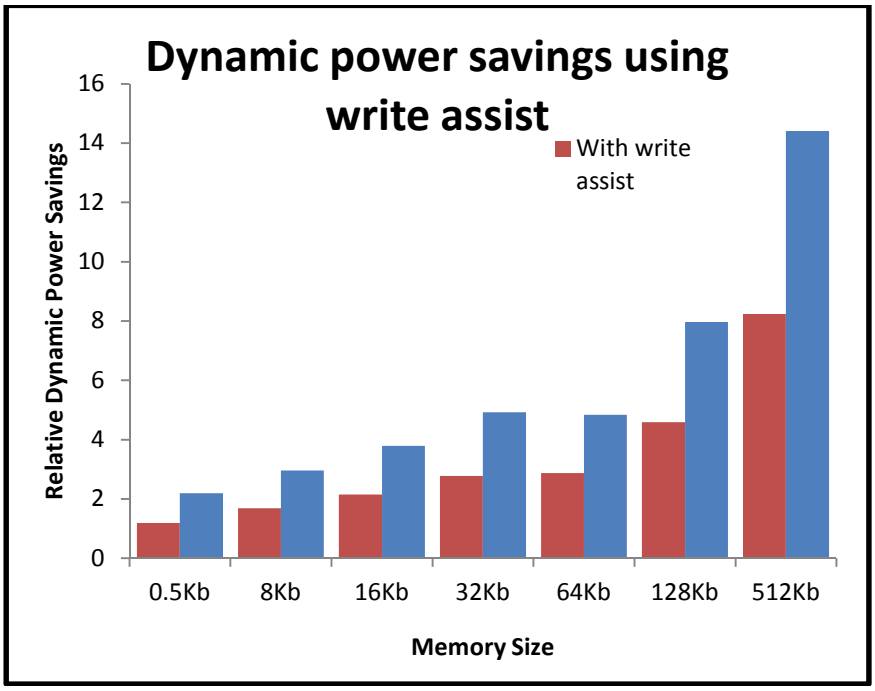

Figure 6: Dynamic power savings using write assist

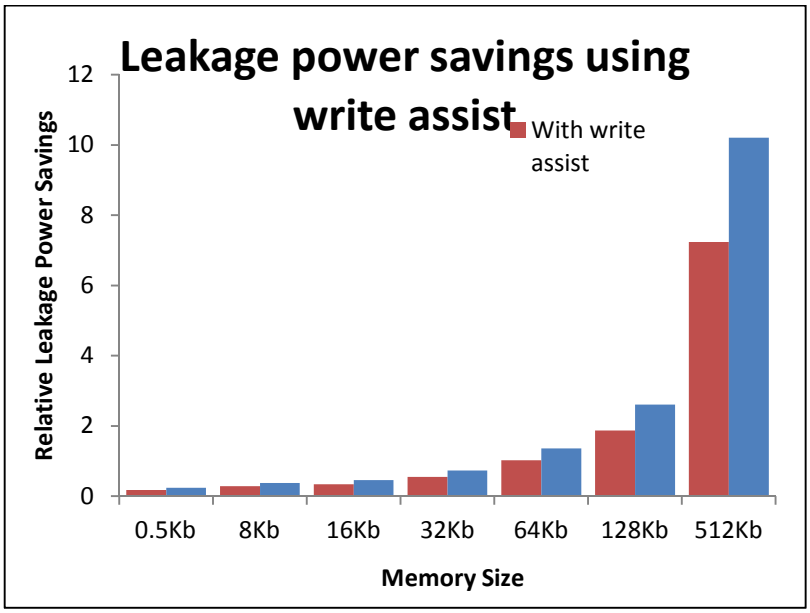

Figure 7: Leakage power savings using write assist

Artisan Low Voltage Compilers Provide Write Assist

The ARM® Artisan® Low Voltage Memory Compilers provides SoC designers with a power optimized memory instances. Artisan Low Voltage Memory compilers are based on a power-efficient architecture that effectively reduces static and dynamic power consumption. Additionally, designers can select high performance or low power options using write assist to further increase the range of applications that can utilize write assist. These memory compilers are ideal for cost effective, low power SoC designs and provide the most effective solution to ensure superior quality. All Artisan Memory Compilers are delivered with a complete set of models for industry-standard EDA tools, correlated with library models, and silicon tested at various voltage and temperature conditions.

Conclusion

As the technology scales, the challenges in designing SRAM arise due to increase in write, read disturb, access and retention failures. It is now becoming common for SRAMs to have read and write assist techniques to enable robust operation at lower supply voltages. In this paper we reviewed the write failure mechanism and write assist techniques. We also benchmarked ARM memories with write assist implementation. For dynamic power we can achieve up to 14% savings with iso-frequency operation and up to 35% savings with DVFS. For leakage power we can achieve up to 25% savings. The ARM® Artisan® Low Voltage Memory Compilers with write assist feature can significantly increase the range of applications that can utilize low voltage operation.

References

[1] V. Chandra, C.Pietrzyk and R. Aitken, On the Efficacy of Write-Assist Techniques in Low Voltage Nanoscale SRAMs, Design Automation and Test in Europe (DATE), 2010.

[2] G. Yeung, G, P. Hoxey, P. Prabhat, Y.K. Chong, D. O'Driscoll and C.Hawkins, Low power memory implementation for a GHz+ Dual Core ARM Cortex A9 processor on a high-K metal gate 32nm low power process, VLSI Design, Automation and Test (VLSI-DAT), 2011.

[3] V. Chandra and R. Aitken, On the impact of gate oxide degradation on SRAM dynamic and static write-ability, Asia and South Pacific Design Automation Conference (ASP-DAC), 2011.

[4] D. Kim, V. Chandra, R. Aitken, D. Blaauw and D. Sylvester, Variation-aware static and dynamic writability analysis for voltage-scaled bit-interleaved 8-T SRAMs, International Symposium on Low Power Electronics and Design (ISLPED), 2011.

[5] N. Shibata et al "A 0.5V 25 MHz 1mW 256kb MTCMOS/SOI SRAM for Solar-Power-Operated Portable Personal Digital Equipment – Sure Write Operation by Using Step-Down Negatively Overdriven Bitline Scheme, IEEE Journal of Solid-State Circuits, Vol. 41, Mar 2006.

[6] US7613052, Memory device and method of operating such a memory device, Van Winkelhoff *et al*.

[7] US8045402, Assisting write operations to data storage cells, Gus Yeung.

[8] C. C. Wang et al, A Boosted Wordline Voltage Generator for Low Voltage Memories, ICECS, 2003.

[9] H. S. Yang et al, Scaling of 32nm Low Power SRAM with High-K Metal Gate, IEEE Intl. Electron Devices Meeting, 2008.

[10] K. Zhang et al, A 3 GHz 70 Mb SRAM in 65nm CMOS technology with integrated column-based dynamic power supply, IEEE Journal of Solid-State Circuits, Vol. 41, No. 1, Jan 2006.

[11] US 8164964, Boosting voltage levels applied to an access control line when accessing storage cells in a memory, Vikas Chandra *et al*.