Reliability Analysis of Power Gated SRAM under Combined Effects of NBTI and PBTI in Nano-Scale CMOS

Anuj Pushkarna
San Francisco State University
1600 Holloway Ave
San Francisco, CA - 94132
Telphone nymber: 1-415-827-6529
anujp@hotmail.com

Hamid Mahmoodi
San Francisco State University
1600 Holloway Ave
San Francisco, CA - 94132
Telephone number: 1-415-338-6579
mahmoodi@sfsu.edu

ABSTRACT
Transistor aging effects (NBTI and PBTI) impact the reliability of SRAM in nano-scale CMOS technologies. In this research, the combined effect of NBTI and PBTI on power gated SRAM is analyzed. Optimal source biasing in the standby mode is presented as an effective method for guard-banding against the aging effects. The simulations results in a predictive 32nm technology shows maximum of 1.6% reduction in standby SNM over 5 year lifetime. For optimum operation, by decreasing the standby source bias voltage by only 0.012 volts, the SNM is safely margined for 5 year life time. This guard-banding comes at an insignificant power overhead of 0.6% for applied worse case scenarios. Given the insignificant power overhead with such guard-banding, it is concluded that adaptive tuning of the source biasing voltage is not required, given the not so negligible complexity and overhead associated with adaptive techniques.

Categories and Subject Descriptors
B.3.4 [**Memory Structures**]: Reliability, Testing, Fault Tolerance

General Terms
Design, Performance, Reliability.

1. INTRODUCTION
Over recent years the scaling in CMOS technology has been very aggressive. With ultra thin dimensions, the technology is facing many critical challenges and reliability issues. One of the very significant issues is caused by transistors, more popularly known as the Bias Temperature Instability (BTI) effects. Evidently there are two types of such effects, Positive and Negative BTI (PBTI & NBTI), happening on NMOS and PMOS transistors respectively. Under NBTI, due to trapping of charges, Vt of PMOS goes up and PMOS becomes slower over time (reduced ON current). Similar issue is the case with PBTI in regard to the NMOS transistor. The manufacturing control gets limited with scaling devices, causing many parameters to undergo some degree of variation. One of the main parameter is the Threshold Voltage (Vth) of the transistors. The variations are generally of two types: Inter-die (die to die) variation and Intra-die (within die) variation. Such variations are caused by many reasons like variations in channel length, channel width; doping, line edge roughness, random dopant fluctuations and oxide thickness. We have used Predictive Technology Model (PTM) of a 32nm CMOS technology [3] and considered nominal Vt conditions in our study. According to the PTM model for 32nm, the value of Vt is taken to be 0.2volts.

Although the power gated SRAM cells have been studied [1], very less amount of work is done on effects of transistor aging on such circuits. Most of the existing work is limited to the study of the impact of NBTI alone. In this paper, we study and report the combined aging effects (NBTI and PBTI) on the conventional power gated SRAM cell. Implementation of power gating introduces a 'Sleep mode' to the cell. Under the standby conditions, the cell becomes cut off from the actual ground supply node. This results in creation of a virtual ground node, or commonly called source bias. We reckon that a power gated SRAM cell can be optimized by applying a particular level of Source Biasing voltage to the virtual ground of the cell. The SRAM design parameters such as Static Noise Margin (SNM), Access time, Write Margin and Leakage power would be used to substantiate the results [5]. We apply the aging effects to the same design and further demonstrate its influence on reliability of the design. The main tasks for this study are as follows:

1) Reporting performance parameters of a Power Gated SRAM Cell.
2) Applying Aging effects (NBTI and PBTI) to the cell and quantifying the impact on different cell performance parameters.
3) Implementing a 64 cell array with a shared Sleep Transistor and obtaining results.
4) Applying aging effects to the 64 cell array and quantifying the impact.

By applying a specific source biased voltage we can choose our operating point on the basis of design requirements. The effect of aging on normal 6T SRAM cell has already been researched [2]. Power gating is relevant for SRAM designs because it is commonly applied for reducing standby leakage power consumption. It is typically applied to larger blocks (rows/columns), which are controlled by standby signals derived by ad-hoc micro-architectural arrangements that use various types of block access statistics.

2. MODEL OF BTI UNDER TEMPERATURE AND PROCESS VARIATIONS
A Predictive model for NBTI/PBTI was recently developed by Predictive Technology Model (PTM) group [6, 4]. We have used this Beta version of the model to estimate the changes in Threshold Voltage Vt over time. The model demonstrates the dependence of NBTI on process and design parameters for particular technology

GLSVLSI'10, May 16–18, 2010, Providence, Rhode Island, USA

used. In our case, the aging model is calibrated for 32nm using High K metal gate model. The aging mainly depends on the mode of operation of the transistor, i.e. static or dynamic. For a particular transistor, with the stress and recovery timing as shown in Fig 1:

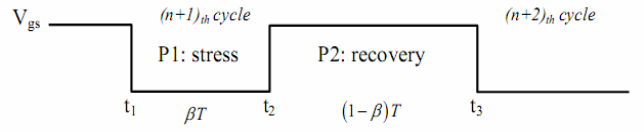

Figure 1: A sample input at the Gate of a transistor under consideration.

Phase 1: Stress, t = (t1, t2)

$$\Delta V_{th} = \sqrt{K_v^{\,2} \cdot (t - t_0)^{0.5} + \Delta V_{th1}^{\,2}} + \delta_v$$

Phase 2: Recovery, t=(t2,t3)

$$\Delta V_{th} = (\Delta V_{th2} - \delta_v) \cdot \left[1 - \sqrt{\eta (t - t_0)/t}\right]$$

$$K_v = A \cdot T_{ox} \cdot \sqrt{C_{ox}(V_{gs} - V_{th})} \cdot \exp\left(\frac{E_{ox}}{E_0}\right) \cdot \left[1 - \frac{V_{ds}}{\alpha (V_{gs} - V_{th})}\right] \cdot \exp\left(-\frac{E_a}{kT}\right)$$

Where, $E_{ox} = (V_{gs} - V_{th})/T_{ox}$

$$\Delta V_{th} = \sqrt{K_v^{\,2} \cdot (t - t_0)^{0.5} + \Delta V_{th1}^{\,2}} + \delta_v$$

For long term degradation and multiple cycle predictions, the following equations describe the Vt change after n cycles of stress and recovery, where β is the stress duty cycle and T is the clock period.

$$\Delta V_{th} = K_v \cdot \beta^{0.25} \cdot T^{0.25} \cdot \left[\frac{1 - \left(1 - \sqrt{\eta(1-\beta)/n}\right)^{2n}}{1 - \left(1 - \sqrt{\eta(1-\beta)/n}\right)^{2}}\right]^{0.5} + \delta_v$$

The default values of the coefficients used in the equations are: A= 1.8 (mV/nm/$C^{0.5}$), α=1.3, E_O= 2.0 (Mv/cm), E_a=0.13(eV), η = 0.35 and δv=5.0(mV). For the 32nm technology that we use, we take Vdd=0.9V (Supply Voltage), Vth= 0.2V and Tox = 1.65nm (Effective Oxide Thickness). The model implemented on takes the values of the following parameters and estimates shift in threshold voltage: Oxide Thickness, Vgs (Gate Stress Voltage), Vth(Threshold voltage), Vds (Drain to source biasing of the transistor), β (Percentage stress time. For DC stress, β=1 and for AC stress, β < 1), Temperature, Life time (for e.g. 2 years or 5 years etc.).

3. ANALYSIS OF AGING EFFECTS IN POWER GATED SRAM

Fig. 2, shows a single cell structure of the power gated SRAM cell that we exercised for our results. It has 2 pull up PMOS (M2, M4 with W=36nm & L=32nm) and 2 pull down NMOS (M1, M3 with W=71nm & L=32nm) transistors, connected in form of two cross coupled inverters and 2 NMOS (M5, M6 W=36nm & L=32nm) access transistors to access the SRAM cell during Read and Write cycles. M7 works as a gate between the cell source line (SS) and the actual ground. If the signal SL is at logic '0', the node SS gets disconnected from the ground. This mode is called as 'sleep mode' for the cell itself. The port SS works as a Virtual ground for the cell, and its logic zero voltage level directly follows the voltage of this node. By applying a positive voltage to the virtual ground node (source biasing), the leakage power of the SRAM array can be effectively minimized. We discuss a method for optimal voltage biasing for this source biased node, acting as virtual ground in Sleep mode.

We can assume Q' to hold a logic '1' and Q being at logic '0'. This figuratively is taken as the initial condition of the cell. For this state, M2 being 'ON' is stressed and experiences the NBTI based Vt shift. Similarly M3 is also 'ON' and due to PBTI

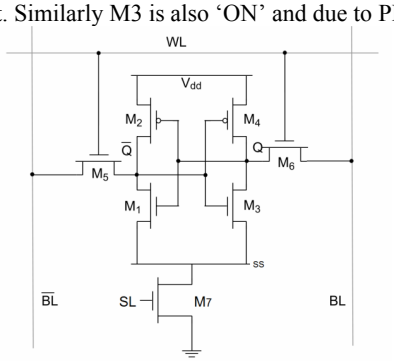

Figure 2: Power gated SRAM Cell. A usual 6T assembly with 1 additional Transistor (Sleep Transistor)

would experience a Vt shift over time. This shift in Vt of these transistors can cause change in performance parameters of the cell. SNM being the most critical parameter degrades over time under the condition of stress as long as the cell holds that value. Under the stress, the trip point for inverter (M1, M2) comes low due to increased Vt of M2. Opposite is the case with M3&M4. This way the cell can very easily get flipped during the read operation. The Write margin degrades after first cycle of write operation as the trip point of the inverter (M1, M2) has reduced. The Access Time of the cell would also increase, as it is determined by discharging process of BL using M3 as the pull down transistor which would be influenced by PBTI. The aging of the sleep transistor (SL) is negligible because for most duration of its lifetime, M7 is in off condition (β≈0). Same is the case with the memory array.

The Sleep Transistor is deployed in a practical design is by sharing a single big transistor between multiple cells per row of the memory array, making it area and cost effective (Figure 3). In the more general situations, a single cell in the whole array of memory can be kept in sleep mode for more than 90% time as compared to the time during which it is accessed for read and write operations. The general degree of efficiency of this scheme for leakage power reduction is known [1] and our experimental results validate this observation as well. Our results tend to justify the minimal changes that tend to happen after applying the aging.

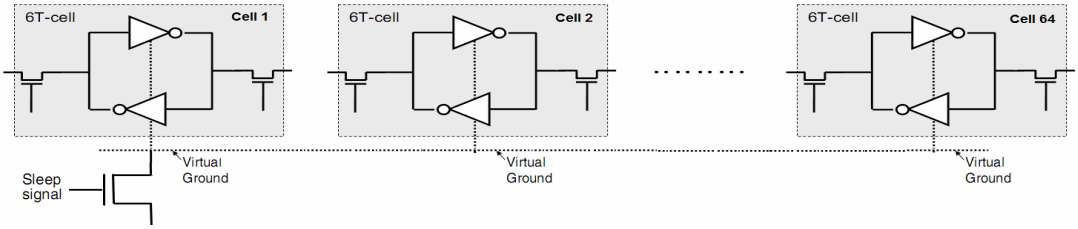

Figure 3: An Array of 64 SRAM cells with a single shared Sleep Transistor.

By simulations we came up with performance results of a single power gated cell as shown in Table 1. These tables show change in performance parameters of a single Power Gated SRAM Cell (Fig. 2).The parameters for consideration are SNM, Access Time and Write Margin. The SNM in this case is measured during the active mode of the cell. This means that the Sleep Transistor M7 is switched 'ON' by applying a logic '1' at the SL port. We can make the following observations from the results of a single power gated cell:

1) SNM (active mode) drops by 2.4% in 25°C and 3.8% in 100°C, for 2 years life time.
2) Access Time increases by 0.26% in 25°C and 0.4% in 100°C, for 2 years life time.
3) Write Margin drops by 2.9% in 25°C and 2.7% in 100°C, for 2 years life time.

The important nature of change in Vt for the stressed transistors was calculated by using the model described in Section 2. It is important to know the value of the parameter 'β' (stress duty cycle parameter for the aging model [4]) to be used for aging estimation of each transistor in the cell. It is certain that if we observe the working of a single cell in a big array of SRAM, it is very unlikely that it would be under stress (of read or write operation) for a substantial time as compared to the whole operation time of the Array itself. Using this approach, the model gives a 5mv (δv in the ΔVth model) change in Vt regardless of the lifetime. Using the same β values and initial conditions, we came up with the resulting change in performance parameters for a 64 cell array of these SRAM cells (Figure 3), as shown in Table 2.

For a 64 cell array:

1) Access time increases by 1.12% in 25oC and 1.14% in 100°C, for 2 years life time.
2) Write Margin drops by approximately 0.5% in 25oC and 1.8% in 100°C, for 2 years life time.

Table 1: Change in Performance parameters for 64 Cells (Array), with BTI effects for 2 Years.

Access Time (nsec)		
	25°C	100°C
Not Aged (0 years)	2.132	2.62
Aged (2 Years)	2.156	2.65

Write Margin (Volts)		
	25°C	100°C
Not Aged (0 years)	0.415	0.433
Aged (2 Years)	0.413	0.425

It would be suitable to compare the performance parameters of a normal 6T SRAM Cell and a power gated SRAM. The model used for the transistors is 32nm High K design. Results with no aging effects are considered. If we compare the Table 1and 3, we can derive the following inferences:

1) The SNM of power gated SRAM cell decreases by only 7.3% when compared to the conventional 6T design. This is not significant drop and could still be very well considered in design envelope.
2) The access time of power gated design reduces by 4.6% when compared to conventional 6T design. This is suitable as we are making the design faster, and is also required.
3) There is no significant difference in write margin of the both the designs, so there is no overhead.

Considering these observations, we can say that the power gated design is very suitable for application to SRAM design. We already know have observed how effective is the power gated design, when it comes to reducing leakage power [1]. A power saving of approximately 80% is very significant.

Table 3: Design parameters of a 1 cell, conventional 6T SRAM

Parameters	25°C	100°C
SNM (Volts)	0.104	0.0701
Access Time (nS)	0.371	0.469
Write Margin (Volts)	0.3366	0.3667

4. OPTIMIZATION OF SOURCE BIASING UNDER AGING EFFECTS

Source biasing can establish a trade-off between the SNM and leakage power in the sleep mode. As the source bias voltage (Vsbv) increases, the SNM of the cell reduces, however the leakage power improves. If we analyze the single cell as in Fig. 2, there seems to be an optimal source bias for minimum leakage because as the source bias increases the leakage of the cell goes down but the leakage of the sleep transistor goes up due to increase Vds. When we consider a large array of sleep transistors with a shared transistor (Fig. 3), the optimal source bias for minimum leakage will be higher as the overall leakage is dominated by the sum of the cell leakages. Hence, in this case, the higher source bias will give more leakage saving, but we are limited by the minimum SNM requirement. It is certain that the Vsbv can only be applied when the SRAM Cell or Array is in sleep mode, i.e. by applying logic '0' at SL. Any voltage that is applied to the port SS, is followed for the cell. Keeping the desired SNM is a chief requirement (to guarantee the reliability of the cell). Meanwhile if we consider the total power dissipation over this sleep mode for both cell and the array, we can come up with an optimal point of operation.

It is also imperative to notice that the SNM (termed SNM_Sleep) would get affected by the aging effects (NBTI and PBTI) over time, because of change in Vt of the transistors that are 'ON'. It is already been speculated that with technology scaling the leakage power plays more critical role. This is good for Power Gating Scheme, as the Aging effects shouldn't influence the Power factor during the Sleep Mode. Based on simulations, we came up with the results of these two parameters (SNM_Sleep and Total Power) for both 1 Cell configuration and 64 Cell configurations as well. The Table 4 (A) shows the change in SNM_Sleep over application of variable Vsbv as well as the aging over time. In regard to the aging effects, we assume $\beta \approx 1$ (parameter from the aging model [4]) for the ON transistors. β is approximated to be zero for the access transistors. This assumption is reasonable considering the fact that cell would be always under stress. Table 4

Table 2: Change in Performance parameters for 1 Cell (Power gated). Both NBTI & PBTI applied for 2 Years.

SNM (Active) - Volts		
	25°C	100°C
Not Aged (0 years)	0.1123	0.078
Aged (2 Years)	0.1096	0.075

Access Time (nsec)		
	25°C	100°C
Not Aged (0 years)	0.389	0.496
Aged (2 Years)	0.39	0.498

Write Margin (Volts)		
	25°C	100°C
Not Aged (0 years)	0.34	0.37
Aged (2 Years)	0.33	0.36

(B) shows the change in total power (nW) over variable Vsbv as well as the aging effects. These measurements are also for the Sleep mode of the 1 cell and 64 cells in an array, for worst case scenario of 100^{O}C.

Table 4: (A) - Readings of 'SNM_Sleep' for variable Source Bias Voltage & Aging effects (can be considered for both 1 & 64cells). (B)- Readings of 'Total Power' (includes Leakage) for variable Source Bias Voltage and Aging effects.

SNM (Stand by Mode – Volts)			
Vsbv (Volts)	Unaged	2 Years	5 Years
0.1	0.21207	0.20898	0.2086
0.15	0.20416	0.20118	0.20091
0.2	0.19452	0.19156	0.19133
0.25	0.18317	0.18011	0.17996
0.3	0.1703	0.1673	0.16713
0.35	0.15622	0.15327	0.15313
0.4	0.1412	0.13832	0.13821
0.45	0.1253	0.1226	0.12251
0.5	0.1088	0.1061	0.10607
0.55	0.09146	0.08892	0.0888
0.6	0.07378	0.07097	0.07093
0.65	0.0554	0.05272	0.05269
0.7	0.03632	0.0338	0.0338
0.75	0.01746	0.0147	0.0147
0.8	0	0	0

(A)

1 Cell at 100^{O}C			
Total Power in Sleep Mode (nW)			
Vsbv (Volts)	Unaged	2 Years	5 Years
0.1	18.632	18.61	18.6
0.15	12.223	12.21	12.2
0.2	8.779	8.76	8.75
0.25	7.2817	7.276	7.27
0.3	7.0115	7.0016	7.001
0.35	8.1323	8.116	8.11
0.4	10.49	10.484	10.48
0.45	13.844	13.834	13.831
0.5	18.65	18.64	18.64
0.55	25.488	25.488	25.488
0.6	34.671	34.671	34.671
0.65	46.902	46.902	46.902
0.7	63.109	63.109	63.109
0.75	84.5	84.5	84.5
0.8	112.65	112.65	112.65

64 Cells at 100^{O}C			
Total Power in Sleep Mode (nW)			
Vsbv (Volts)	Unaged	2 Years	5 Years
0.1	2015.1	2013.8	2013.7
0.15	1611.6	1620.01	1620.01
0.2	1367.5	1366.7	1366.7
0.25	1211.3	1210.6	1210.5
0.3	1115.06	1118.8	1118.7
0.35	1066.5	1065.4	1065.4
0.4	1032.8	1033.9	1033.8
0.45	1025.3	1020.1	1020
0.5	1013.8	1008.01	1007.9
0.55	995.5	986.53	986.36
0.6	957.82	947.45	947.28
0.65	921.67	915.24	915.13
0.7	906.08	893.75	893.75
0.75	884.31	868.88	868.88
0.8	850.7	828.86	828.86

(B)

Based on the results in Table 4, we plot the in order to identify the optimum point (Fig 4). The SNM in sleep mode drops proportionally as we increase the source bias voltage Vsbv. The same curve can be used to evaluate SNM for both 1 Cell and 64 Cells scenarios. We can easily observe an optimum bias point. This comes out to be for Vsbv=0.3 V, where the SNM (Sleep) =0.17V. These results approximately remain same even after applying aging effects for 2 and 5 years. The maximum drop in the SNM is observed to be 1.6% for operating points corresponding to Vsbv=0.1V. As observed from curve this difference drops down to almost negligible. In the more practical application in the form of 64 Cell array, the total power vs. source bias voltage (Fig. 4) does not shows any local minima. The minimum requirement on SNM (Sleep mode) would be the deciding factor to select the maximum (optimal) source bias voltage. It is tangible from the nature of the curve (for 64 Cells), that there are no significant changes in total leakage power for the points around Vsbv=0.5V (the leakage curve is almost constant around this bias point).

It was speculated that in order to maintain the SNM (sleep) over time, there would be a need to calibrate Vsbv (For 64 Cell array). However the results clearly show that such process of Vsbv calibration could be avoided, as there is negligible shift in SNM in sleep mode. Instead, it is much easier to provide margin for SNM by lowering Vsbv during design in order to tolerate the aging effects. For e.g. in order to maintain SNM of 0.12V over a period of 5 years, we can lower Vsbv to ≈ 0.4879V (compared to 0.5V if there were no aging effects). This slight reduction in Vsbv voltage provides enough SNM margin against the aging effects and comes at the insignificant power overhead of 0.6%.

5. CONCLUSION

We analyzed the reliability of power gated SRAM under combined NBTI and PBTI aging effects in nano-scale CMOS technology. We have shown that power gating of SRAM can significantly minimize the instability that the aging effects might cause. Moreover, our results show that sufficient SNM margin against aging effects can be established by lowering the source bias voltage in the sleep mode and this comes at a negligible power overhead.

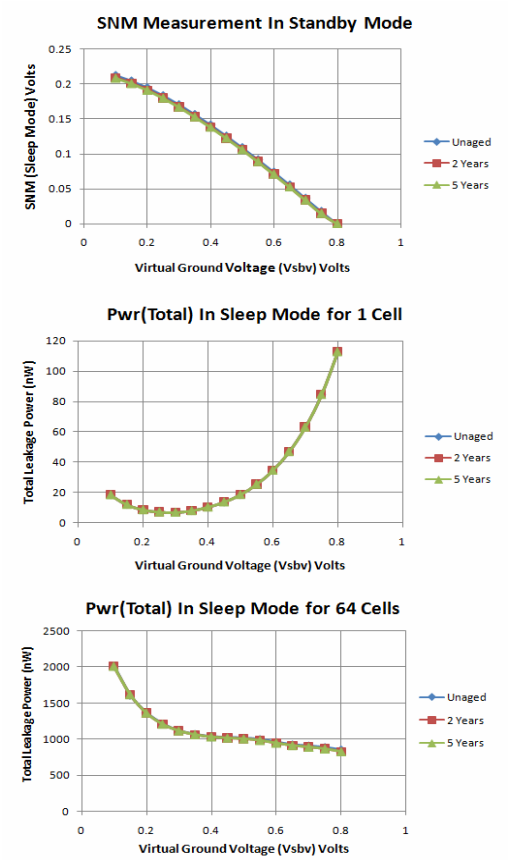

Figure 4: Graphical representation for Tables 3&4. SNM_Sleep & Total Power for both 1 and 64 Cell configurations.

The results show a maximum of 1.6% reductions in standby SNM over 5 year lifetime. For optimum co-ordinates by decreasing the standby source bias voltage by 0.012 volts, the SNM is safely margined for 5 year life time. This reduction in the source bias voltage comes at an insignificant power overhead of 0.6%. More importantly we also derive an optimum performance co-ordinate for the design by selecting a particular source bias voltage, along with a reasonable way for selecting it.

6. REFERENCES

[1]. Zhang, Bhattacharya, Chen; "SRAM Design on 65nm CMOS with Dynamic Sleep Transistor for Leakage Reduction", Solid-State Circuits, IEEE Journal of Publication, Date: April 2005 Volume: 40,Issue: 4 pp: 895- 901

[2]. K. Kang, et. al., "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," IEEE Transactions on Computer Aided Design of Integrated Circuits and System, vol. 26, no. 10, pp. 1770-1781, Oct. 2007

[3]. Predictive Technology Models (PTM): http://www.eas.asu.edu/~ptm/

[4]. PTM Reliability models (NBTI) http://www.eas.asu.edu/~ptm/

[5]. S. Mukhopadhyay, et. al.; "Modeling of Failure Probability and Statistical Design of SRAM Array for Yield Enhancement in Nanoscaled CMOS", IEEE Vol. 24, No. 12, December 2005.

[6]. R. Vattikonda, W. Wang, Y. Cao, "Modeling and minimization of PMOS NBTI effect for robust nanometer design," DAC, pp. 1047-1052, 2006.

[7]. Zhang, Guo; "Active Leakage Control with Sleep Transistors and Body Bias". Symposium on Low Power Electronics and Design, 2005.