Robust flip-flop Redesign for Violation Minimization Considering Hot Carrier Injection (HCI) and Negative Bias Temperature Instability (NBTI)

Naeun Zang and Juho Kim

Computer Science and Engineering, Sogang University
Seoul, 121-742, Republic of Korea
spikezang@gmail.com

Computer Science and Engineering, Sogang University
Seoul, 121-742, Republic of Korea
jhkim@sogang.ac.kr

Abstract
As the CMOS device becomes smaller, the process and aging variations become one of the major issues for circuit reliability and yield. Thus, a number of studies on the aging effects are currently underway. In this paper, we measure the setup/hold time and the variations considering aging effects such as a hot carrier injection (HCI) and negative bias temperature instability (NBTI) on flip-flop. The measured data was applied to the transistor sizing algorithm. We also have applied aging effects for 5 years with setup time variation reduction to redesign a more robust flip flop. The proposed method analyzed aging effects (NBTI, HCI) for flip flop at the transistor level in 45nm process and used PTM (predictive technology model) SPICE model. The redesigned flip-flop using the proposed algorithm confirmed to have violation minimization after 5 years.

Keywords: CMOS, Reliability, Aging, NBTI, HCI, Flip-flop

1. Introduction
Lately, miniaturization of semiconductor process to a nanometer unit are causing the problem of process variation and aging with the rapidly increasing circuit density. Process variation is the variation that appears in the device parameters that affect circuit performance such as oxide film thickness (TOX), length of the effective channel (LEFH), width of the effective channel (WEFF), and threshold voltage (VTH). Process variations occur at each process step in semiconductor production, bringing about differences between the actual device parameters and the intended device parameters initially planned at the design phase. Aging phenomenon lengthens the delay time of the circuit by increasing threshold voltage and reducing carrier mobility over time. Aging produces degradation and malfunction of the circuit and this situation could create serious problems. For example, if important information were stored on the flip-flops incorrectly due to the aging phenomenon, it would arise a serious problem about reliability. Thus, at the design stage, the impacts of aging on the circuit performance should be analyzed. Since the accuracy of the analysis is linked to the reliability of the circuit, the circuit needs to be more accurately analyzed herein.

In this paper, we will examine the impacts of aging on setup/hold time of the flip-flop in detail. Unlike the conventional probabilistic analysis method, in particular, we analyzed the aging of transistors that affect the setup/hold time using the stress time of each transistor. From the analysis results, we concluded that hold time can be ignored and therefore only the setup time was considered in applying the sizing algorithm. The redesigned flip-flop minimizes the setup time variation even with aging effect. In addition, the proposed solution minimizes the violations within the minimized range of the power increase.

This paper is organized as follows. In section 2, we described the concept of HCI, NBTI and the threshold voltage variation model, circuit aging analysis method at the transistor level, and the characteristics of the flip-flop. In section 3, we analyzed the impact of aging on the setup/hold time by applying aging phenomenon to the transistors in the flip-flop. In Section 4, we compared the setup/hold time among the following three cases and measured the variations: the original flip-flop, flip-flop redesigned using
the proposed method, and flip-flop after aging is applied. In Section 5, we ended with the conclusion and future agenda to challenge.

2. Aging Modeling.

2.1 Aging Phenomenon

In gate and circuit as a collection of transistors, the delay time is determined by the influence of the transistor characteristics such as the oxide film thickness and threshold voltage. Just as things are aging over time, the transistor is also aging, which gradually change the size of the characteristics. Eventually, due to the aging of the transistor, the delay time of the circuit is also gradually changed differently from what it was at design stage in duration, which will act as a factor to generate a circuit malfunction after a period of time.

There are two typical aging phenomena in transistor: NBTI that occurs in PMOSFET, and HCI that occurs in PMOSFET and NMOSFET. [1][2]

2.1.1 NBTI (Negative Bias Temperature Instability)

NBTI is an aging phenomena occurring in PMOS. When PMOS gate-source voltage (Vgs) is less than 0 (stress interval), the interface trap is created with the separation of Si-H bonds between the oxide film and the substrate by an electric field arising in the oxide film. Finally, due to the creation of interface traps, PMOS threshold voltage would be gradually increased over time. Fig. 1 shows the mechanism of NBTI [6].

In the recovery interval, when gate-source voltage of stressed devices becomes zero again, H comes back to the interface between the oxide film and the substrate, and Si-H bonds are recombined, reducing a portion of the increased threshold voltage. That is, as seen in Fig. 2, NBTI phenomenon shows the stress interval that the threshold voltage is increasing and the recovery interval that some of the increased threshold voltage decrease again.

As a result, the static NBTI analysis, which does not consider the actual behavior of PMOS, will bring about the pessimistic results. More accurate analysis of NBTI is performed in recent studies for NBTI which takes into account the recovery effects with utilization of the duty cycle having the meaning of ON / OFF ratio[3][4].

The following is the equation of the long-term prediction model (1) presented by W.Wang, et al., where the variations in the threshold voltage are approximated by the utilization (α) and the input clock period (Tclk). Variation in threshold voltage for each flip-flop transistor was calculated using the above equation.

The meaning of each parameter is described in the reference [3].

\[ \Delta V_{th_{NBTI}}(t) = \left( \frac{\sqrt{K_{ox}^2 \alpha T_{clk}}}{1 - \beta(t)^{1/2n}} \right)^{2n} \]  

\[ \beta(t) = 1 - \frac{2\xi_1 \cdot t_c + \sqrt{\xi_2 \cdot C \cdot (1 - \alpha) \cdot T_{clk}}}{2t_{ox} + \sqrt{C \cdot t}} \]
2.1.2 Hot Carrier Injection

As with BTI, HCI is created by generation of the interface state between the silicon in the MOSFET and the oxide film. However, the causes that the interface states arise differ between the two aging phenomena. Fig. 3 depicts the cause of HCI. HCI arise as the hot carriers flow into the gate oxide while hot carrier is generated when carriers, like a hole in the NMOS and PMOS, move the channel formed between the source and the drain. Carriers such as the electron of NMOS or the holes of PMOS move along the channel formed between the source and the drain. Some carriers become hot carriers, when they obtained a sufficient kinetic energy enough to flow into the gate oxide film away from the channel. By absorbing the electric fields generated in the gate oxide along channel, carriers obtain such a kinetic energy, most of which flows from the drain of the MOS. Due to these phenomena, the boundary surface between the oxide film and the channel state is changed, which results in a change in device parameters such as threshold voltage over time.

\[
\Delta V_{th_{HCI}}(t) = \left( \frac{I_d}{I_{d,ref}} \right)^{m} t^n
\]  

(3)

\( I_d \) and \( I_{sub} \) represent the drain current and the substrate current of the transistor, respectively. 

\( W \) is the width of the transistor. \( H \) and \( m \) are voltage acceleration coefficients. \( t \) is a total time stressed and \( n \) refers to the time acceleration coefficient.

2.2 Flip-Flop

Flip-flop is a memory element to store one bit of information, referring to a device or circuit which can temporarily hold or store the state of the signal, and change the current binary information depending on the input and state given by the pulse.

Flip-flop can store one bit of information since it alters into the opposite state from the current state (from 0 to 1, or from 1 to 0) when a voltage is applied, sustaining the status on and on. In addition, the point in time at which the output reflects the input is determined in the moment edge of the clock signal. It is composed of a number of transistors and used to configure the SRAM or hardware registers. The types of the flip-flop are diverse, including RS flip-flop, D flip-flop, JK flip-flop, etc. Fig. 4 shows the combined circuit comprising a flip-flop.

Fig. 4 The combined circuit comprising a flip-flop.

2.2.1 Setup/Hold time

The time taken to reach from the current flip-flop to the next flip-flop should always be within one clock cycle. The flip-flop has two limitations in time. One is set-up time and the other is hold time. Set-up time means the time taken for the result value to reach before the next clock starts.

That is, it refers to the minimum holding time required for the input to be correctly recognized before switching takes place. When the set-up time is aligned exactly to one clock, its value may be stored normally in the next clock, but may not be in some case, so it should always be set to have a
certain margin to some extent.

When the design is made on a block-by-block basis, if flip-flop is not designed to be placed directly inside, but is designed to go through some combinational circuit, a larger set-up time may be required.

Hold time means the time period required for signal to be maintained while the value is stored in the flip-flop. That is, it refers to the minimum time necessary for the change of state to be correctly recognized after switching occurred, which means the minimum amount of time required for the identified result to be maintained.

2.2.2 Setup/Hold time

The set-up violation is the problems, occurring when the setup time or hold time is not met each other, which may arise due to lack of buffers in synthesis constraints. For example, assuming that a hardware structure is capable to work up to 100MHz using 0.5um process, if the designer attempted to design and synthesize the hardware structure to be operated up to 133MHz, a number of set up violation would occurs in the resulting composite. As seen in Fig. 5, a setup violation occurs if the input value A transported to the output Q has not yet been determined after the clock of the flip-flop reach a rising edge. If it is assumed that setup time requirement of a circuit is 1ns and the clock cycle is 10ns, it means that the valid input data A should reach 1ns before the flip-flop clock reach the rising edge as shown in Fig. 6.

The hold time requirement is present in all the sequential logic and, valid data should continue to stay unchanged during the hold time. If valid data doesn’t stay during the hold time, the hold time violation occurs. It usually occurs when designed too rapidly. For example, hold time violation occurs if the input value alters before flip-flop at the rising edge even recognizes the input value.

3. The aging-resistant flip-flop redesign process

In this paper, the aging-resistant flip-flop redesign process is largely divided into two steps. First, we have measured the variations in threshold voltage of the transistors inside of the flip-flop when exposed aging effects (NBTI, HCI) for 5 years and the setup and hold time of the flip-flop[15].

<table>
<thead>
<tr>
<th>DFFS</th>
<th>Stress time (5years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M305fr</td>
<td>69789168s 0.0469V</td>
</tr>
<tr>
<td>M312fr</td>
<td>77231664s 0.0472V</td>
</tr>
<tr>
<td>M319fr</td>
<td>77870248s 0.0472V</td>
</tr>
<tr>
<td>M324fr</td>
<td>78735366.2s 0.0472V</td>
</tr>
<tr>
<td>M331fr</td>
<td>77239548s 0.0472V</td>
</tr>
<tr>
<td>M335fr</td>
<td>38990716s 0.0423V</td>
</tr>
<tr>
<td>M341fr</td>
<td>0s 0V</td>
</tr>
<tr>
<td>M347fr</td>
<td>60257806.2s 0.0462V</td>
</tr>
<tr>
<td>M355fr</td>
<td>46702054.6s 0.0426V</td>
</tr>
<tr>
<td>M362fr</td>
<td>0s 0V</td>
</tr>
<tr>
<td>M366fr</td>
<td>78064214.4s 0.0473V</td>
</tr>
<tr>
<td>M374fr</td>
<td>78064214.4s 0.0473V</td>
</tr>
<tr>
<td>M378fr</td>
<td>40507929s 0.0424V</td>
</tr>
<tr>
<td>M384fr</td>
<td>73202545.8s 0.047V</td>
</tr>
<tr>
<td>M389fr</td>
<td>72239548s 0.0472V</td>
</tr>
<tr>
<td>M395fr</td>
<td>40507992s 0.0424V</td>
</tr>
<tr>
<td>M402fr</td>
<td>39988042.2s 0.0423V</td>
</tr>
</tbody>
</table>

Fig. 7 is the one that structurally depict variations in threshold voltage of transistors inside of the flip-flop caused by the stress that imposed on the transistors for 5 year’s aging periods.

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.
With the setup/hold time measurements, we could conclude that the variations in the threshold voltage affects only the setup time (so the hold time can be ignored). Therefore, we have applied the sizing techniques taking into consideration only the setup time.

In the second step, the variation amounts were arranged in the descending order based on the variation amounts in threshold voltage obtained in the first phase. Lots of efforts have been made to get the post-aging set-up time as close as possible to the pre-aging setup time of the flip-flop by applying the sizing algorithm. The flip-flop can be redesigned to tolerate the stress which is imposed for 5 years without set-up violation using this method. Fig. 8 shows the overall flowchart of a method that is proposed.

3.1 Setup/Hold time measurement

The conventional probabilistic analysis techniques are to transmit the probability of the first input signal to the final output step by step. This method can compute the probability for the output easily by applying the probability of the same input equally to all the transistors in the configuration for the series configuration of the transistors. However, in the case of the parallel structure of the transistor, it is more complex to calculate the probability of the corresponding output and to derive the results, so it may result in more ambiguities in applying the result probability to the feedback inputs in the flip-flop. In addition, it is difficult to measure the exact degree of stress due to the difficulties in reflecting the delay time while the probability of the signal is being transmitted. Conventional probabilistic analysis techniques have drawbacks that it is difficult to consider the transmission delay time inside of the flip-flop and to compute the exact stress level at the transmission gate or the pull-up or pull-down [8][10].

The proposed method analyzed aging of the flip-flop accurately to complement these drawbacks. For analysis purpose, each of the output waveform was analyzed for the four input waveforms that can happen. Status of the four input waveform is classified into the following 4 categories as shown in the Fig. 9: the low to high, the low to low, the high to low, and the high to high. And the stress time was measured for each waveform. The stress times of NBTI and HCI were computed using different model each other when analyzing waveforms. Stress-time linking to the row signal is applied to a PMOS affects NBTI [9].

Thus, only the time corresponding to the low signal, excluding high signal and the transition portion (signals 0.3 to 0.7), is regarded as a stress time. HCI is affected by the transition time that arises in NMOS. So the transition time should be measured and included in the stress time in NMOS. The measurement should be conducted for one clock as a unit time in measuring a stress time. Stress time per clock that the actual transistor is exposed to can be obtained by applying switching activity and the signal probability to the stress-time linked to the respective input state, where the switching activity is the probability of the input signal.
transition, while the signal probability is a probability that
the input come in high.

Fig. 9 shows the stress-time and stress-time of the NBTI
HCI according to the respective input states. Each stress time
was computed for the NBTI and HCI obtained above and the
long-term prediction model and Lucky electron model were
applied to them respectively, and then the changed threshold
time of the final aged transistors was computed as well.
And out of all the transistors of the flip-flop, the transistor
that has been affected the most is identified [11].

The way to obtain the stress time is as follows. Assuming,
for example, that switching activity = 0.7, signal probability
= 0.6, stress-time according to the shape of each input can be
calculated as follows: [12] [13] [14]

Stress time(Low->High) * 0.7 * 0.4 = a
Stress time(High->Low) * 0.7 * 0.6 = b
Stress time(Low->Low) * 0.3 * 0.4 = c
Stress time(High->High) * 0.3 * 0.6 = d

The stress-time of the transistor is the sum of the four
values that are obtained by individually multiplying the
switching activity and signal probability to the stress time
per unit clock for the four inputs. So stress time in the flip-
flop for one clock is the sum of all the values a, b, c, d.

The input data is changed by the aging process for set-up
time. When the input data is changed,

As shown in the figure above, the unwanted input value is
generated. This phenomenon is called the set-up time
violation. If a violation occurs, it becomes difficult to obtain
the credible and accurate measures, causing the problem
with reliability. However, the aging progress doesn’t
generate the violation for the hold time. Fig. 10 shows the
hold time changes with aging. In this study, the experiment
was conducted considering only the set-up time, not the hold
violation.

![Fig. 9 Conditions for computing the full stress time in flip-flop.](image)

![Fig. 10 Hold time changing with aging.](image)

3.2 Reduced variation in the set-up time through the
transistor sizing

The proposed sizing algorithm, when applying aging of 5
years to a flip-flop, conduct sizing in the same descending
order among the transistors inside of the flip-flop that
arranged in the descending order with respect to the
threshold voltage. As shown in Fig. 11, it can be seen that
M_319_ff has a highest threshold voltage.

![Fig. 11 Transistor sizing greatly affected by aging.](image)
started with sizing algorithm applied.

When executing sizing, the set-up time of aged transistor is decreasing and getting closer to the set-up time of the transistor without aging where the sized value can be adopted. When the sizing of the first transistor is completed, the sized transistor becomes the new basis for the next sizing. Next, the sizing is conducted for M_366_ff which has the second highest threshold voltage. M_366_ff will be compared with M_319_ff after sizing and if M_366_ff is less than M_319_ff in set-up time, the sizing process will continue. The set-up time of M_319_ff will also be compared with that of the transistor where the aging is not applied and if set-up time of M_319_ff is less than that of the transistor where the aging is not applied, the sizing process will end. If the set-up time is larger than that of the transistor where sizing started first, sizing doesn’t continue any longer. Next, sizing is conducted with regard to the transistor having higher threshold voltage, where the sizing process shall be repeated until the set-up time with aging is getting closer to the existing set-up time without aging. That is, the aging-resistant robust flip-flop can be redesigned, where the setup time, even after aging is applied, gets closer to the existing set-up time before aging is applied. Such a method is schematized in Fig. 11 above.

---

Fig. 12 Sizing algorithm.

Fig. 12 shows the sizing algorithm. In the below algorithm, ‘setup aged’ is the variable to indicate when the set-up time is aged, and ‘queue’ arranges the threshold voltages of the transistors measured after 5 years’ aging in the descending order. ‘target_transistor’ refers to the transistor having the highest threshold voltages of all transistors in the queue, which size ‘target transistor’. ‘setup_original’ refers to the original set-up time where the aging process are not applied and ‘setup aged’ refers to the set-up time when the aging process is applied for five years. ‘aged_setuptime_sizing’ is the set-up time of the aged flip-flop where the sizing has been performed.

---

Fig. 13  Overall flowchart.

Fig. 13 illustrates the overall flow chart where the amount of change in the set-up time is reduced by expanding the size of the transistor in consideration of the threshold voltage variation of the total transistors in the flip-flop. Here, ‘Tri’ refers to the i-th transistor, while ‘k’ refers to the total number of transistors inside of the flip-flop. Next, ‘optimize value’ means the optimized set-up time that is measured after expanding size, which is overwritten until all the sizing processes of all transistors end. Finally ‘sized setup time’
refers the set-up time after sizing is completed. In the algorithm, the amounts of changes in threshold voltage as measured in the transistors are arranged in the descending order while the changes in setup time due to the aging are reduced by expanding the transistor with the largest amount of change in the threshold voltage among the arranged transistors.

4. The experimental results and analysis

In order to accurately analyze and validate the aging phenomenon in the flip-flop, the proposed method in this paper applied the temperature of 25 °C and the supply voltage of 0.9V in 45nm process, using DFF, DFFS, DFFR, DFFRS, SDFF, SDFFS SDFFR, and SDFFRS. Predictive technology model (PTM) was used as SPICE model. And long-term prediction model was used to measure the variation in the threshold voltage for NBTI while Lucky electron model was used for HCI.

Table 2 shows variations in setup/hold time and the variation in the threshold voltage both when aging is applied and when aging is not applied. As shown in Table 2 shows how set-up time is increased when aging is in progress indicating that it is the direct cause of the setup violation. On the other hand, it also indicates how hold time is decreased when aging is in progress, and it allows to confirm that hold violation do not occur.

The transistors are arranged in the descending order with regard to the size of the variation in the threshold voltage when aging is applied for 5 years. When 10% upsizing is performed in the arranged order, both the setup time before aging and after aging are reduced, as identified in table 3.

When applying sizing algorithm based on the experiment results, the set-up time after aging is getting closer to the setup time not affected by aging, which minimize the probability of violation.

Table 2: Variations in setup/hold time of DFSS

<table>
<thead>
<tr>
<th>Flip-Flop</th>
<th>Fresh</th>
<th>Hold</th>
<th>△Setup/△Hold</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFFS</td>
<td>1.80E-11</td>
<td>1.90E-11</td>
<td>5.56%</td>
</tr>
<tr>
<td></td>
<td>-2.30E-11</td>
<td>-2.70E-11</td>
<td>-17.39%</td>
</tr>
<tr>
<td>DFF</td>
<td>2.00E-11</td>
<td>2.00E-11</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>-1.10E-11</td>
<td>-1.40E-11</td>
<td>-27.27%</td>
</tr>
<tr>
<td>DFFR</td>
<td>2.50E-11</td>
<td>2.53E-11</td>
<td>1.20%</td>
</tr>
<tr>
<td></td>
<td>-0.90E-11</td>
<td>-1.46E-11</td>
<td>-62.22%</td>
</tr>
<tr>
<td>DFFRS</td>
<td>2.50E-11</td>
<td>2.53E-11</td>
<td>5.56%</td>
</tr>
<tr>
<td></td>
<td>-0.90E-11</td>
<td>-1.46E-11</td>
<td>-17.39%</td>
</tr>
<tr>
<td>SDFF</td>
<td>5.00E-11</td>
<td>5.30E-11</td>
<td>6.00%</td>
</tr>
<tr>
<td></td>
<td>-5.20E-11</td>
<td>-5.20E-11</td>
<td>0.00%</td>
</tr>
<tr>
<td>SDFFR</td>
<td>5.20E-11</td>
<td>5.40E-11</td>
<td>3.85%</td>
</tr>
<tr>
<td></td>
<td>-5.30E-11</td>
<td>-5.30E-11</td>
<td>-0.42%</td>
</tr>
<tr>
<td>SDFFRS</td>
<td>1.49E-11</td>
<td>1.58E-11</td>
<td>6.04%</td>
</tr>
<tr>
<td></td>
<td>-5.40E-11</td>
<td>-5.40E-11</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 3: Variations in setup time with sizing

<table>
<thead>
<tr>
<th>Flip-Flop</th>
<th>Setup time (sec)</th>
<th>Hold time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-sizing</td>
<td>2.50E-11</td>
<td>3.00E-11</td>
</tr>
<tr>
<td>m324</td>
<td>2.40E-11</td>
<td>2.80E-11</td>
</tr>
<tr>
<td>m324, m366</td>
<td>2.40E-11</td>
<td>2.70E-11</td>
</tr>
<tr>
<td>m324, m366, m374</td>
<td>2.40E-11</td>
<td>2.70E-11</td>
</tr>
<tr>
<td>m324, m366, m374, m319</td>
<td>2.38E-11</td>
<td>2.60E-11</td>
</tr>
<tr>
<td>m324, m366, m374, m319, m384</td>
<td>2.30E-11</td>
<td>2.60E-11</td>
</tr>
<tr>
<td>m324, m366, m374, m319, m384, m389</td>
<td>2.20E-11</td>
<td>2.48E-11</td>
</tr>
</tbody>
</table>

5. Conclusion

This study redesigned a flip-flop with more aging tolerance. The impacts of aging on setup/hold time were conducted applying 5 year’s aging to DFFR, DFFS, DFFRS, SDFF, SDDF, AND SDFFR. From the experiment results, we have concluded that the hold time violation can be ignored in the flip-flop where aging is in progress. In addition, we have proposed a flop-flop that can minimize the setup violation by applying transistor sizing algorithm.

References


Naeun Jang received a B.S. degree from the Department of Computer Science and Engineering at Sogang University in 2002. He received an M.S. degree in Computer Science and Engineering from Sogang University in 2004. He is currently pursuing a Ph.D. degree in Computer Science and Engineering at Sogang University. His research interests include Low Power and High Performance Design Methodology, Statistical Timing Analysis, Cell Characterization, Crosstalk, Aging and Yield analysis.

Juho Kim received the B.S. and Ph.D. degrees in computer and information science from the University of Minnesota, Minneapolis, in 1987 and 1995, respectively. He was a Senior Member of Technical Staff with Cadence Design System, San Jose, CA, in 1997. He joined the Department of Computer Science and Engineering, Sogang University, Seoul, Korea, in 1997, and was a Department Chair from 2005 to 2008.