ISSN 0002-306X. Изв. НАН РА и ГИУА. Сер. ТН. 2012. Т. LXV, № 1.

UDC 621.382.13 RADIOELECTRONICS

V.SH. MELIKYAN, H.P. PETROSYAN, A.M. DURGARYAN, E.H. BABAYAN
N.KH. ASLANYAN

RETENTION FLOP BOOSTING MECHANISM FOR SELF SAVE/RESTORE CAPABILITY

Power gating has been widely employed to reduce sub-threshold leakage. Data retention flops (RF) are used to preserve circuit states during power down, if the states are needed again after wake-up. These elements must be controlled by an extra power management unit, causing a network of control signals implemented with extra wires and buffers. In this paper analytical expression for RF parameters are presented. The dependency of those parameters from transistor sizes, technological factors, node potential values are shown. New mechanism for accelerating RF self save and restore capability is suggested. Proposed boosting mechanism decreases the charging time of virtual ground rails. This helps retention elements avoid short-circuit current while making transition to standby mode and drastically increases the self save and restore time of RF. Experiments were performed on various benchmark circuits taken from ISCAS family, which were synthesized by 45-*nm* technology. The results show that the proposed boosting mechanism increases the speed of self saving and restoring operation, as a result this brings to 32.6% cut in static power consumption.

Keywords: power gating, retention flop, static power, save time, restore time, leakage current.

Introduction. MOSFET scaling into deep sub-micron region has resulted in significant increase in leakage power consumption. Particularly, in 45-*nm* technology and beyond, leakage power consumption will catch up with, and may even dominate, dynamic power consumption [1]. Power gating technique is the most popular technique to suppress the leakage power consumption [2]. The idea of this technique is to cut off the circuit from its power supply rails during standby mode. Due to the supply power switched off, the circuit state, which is represented by logic values in sequential elements of the circuit, is lost. Therefore a need to restore circuit state after power down arises.

The main approach to restore the circuit state from power down mode relies on a dedicated circuit element for state retention. Such an element is flip-flop which is capable of retaining a state during power gating; it is called a retention flip-flop (RF) [3]. In this approach some part of the circuit flops are replaced with the RF. As a result the RFs remember their values during power down mode and restore the circuit state after power activation.

There are various implementations of RFs [4]. However in all the schemes the working frequency of RF is lower than the normal flop frequency. All these flops have two externally controlled inputs by means of which the saving and restoring of the flop data is performed. For controlling these RF new inputs control logic should be implemented and all flop inputs should be connected to that controller. Because of such connection the

whole chip wiring length will increase. The analyses show [5] that the wire length of power gated sequential circuits can be increased by 29% to 60%. As a result of such wiring increase, the dynamic power consumption of wires will increase drastically, and even exceed the power saved during power gating.

General parameters of retention flop. The first parameter, which describes the RF operation before entering into power down mode, is data saving time [6]. The data saving time defines the amount of time which is needed for RF to remember the master flop data. In other words data saving time is the period of time that needs to be passed, for retention logic to remember the master flop data after activation of RF saving input. The equation of data saving time can be written as

$$T_{dsv} = T_{write}^{ret\_logic} - T_{sa}, \quad (1)$$

where T_{dsv} is flop data saving time, $T_{write}^{ret\_logic}$ is the time when the main flop data is written in retention logic and T_{sa} is the flop saving input activation time. The expression for data saving time will be calculated based in Fig.1. The Fig. presents the master latch and retention logic part of the balloon RF [4]. Since nearly all RFs use similar structure as in balloon type RF, the presented calculation scheme can be used for all type of RFs, with slight modification.

Usually RF flops consist of three parts, master, slave and retention latches. The retention latch is the part of RF which is responsible for keeping the retention data in power down mode [4]. The left part of Fig. 1 is the master latch of the RF, the right part of the flop which is prepared for saving the retention data during power down. That part of the RF will be assumed as retention slave latch, in short retention latch, because it is supposed to keep the retention data during power down. As the purpose is to find the expression for flop data saving time, hence it will be assumed that master latch and retention latch have different logic values, for example it is assumed that the A point voltage is equal to supply voltage VDD and the B point voltage is equal to ground voltage VSS.

To satisfy such assumption the M_{P1}^{M} PMOS transistor should be ON in triode region and the M_{N1}^{M} NMOS transistor should be in OFF state for keeping the value of point A equal to VDD. For retention logic to keep the value of point B equal to ground voltage the M_{P2}^{R} PMOS transistor should be in OFF state and M_{N2}^{R} NMOS transistor should be ON in triode state. As a current equation for triode region, the expression given in [7] will be used:

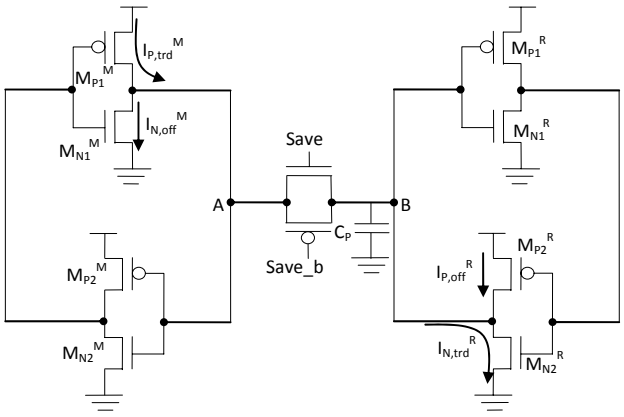

Fig. 1. Master and retention slave latch of RF

$$I_D = \mu C_{OX} \frac{W}{L}((V_{GS} - V_{TH}) * V_{DS} - \frac{1}{2} V_{DS}^{2}), \quad (2)$$

where I_D is the transistor current in triode region, μ is the mobility of channel charge curriers, C_{OX} represent the total capacitance per unit length, W is transistor channel width, L is transistor channel length, V_{GS} is gate-source voltage, V_{TH} threshold voltage and V_{DS} is drain-source voltage. To have more accurate results, the OFF currents of M_{N1}^{M} and M_{P2}^{R} transistors will be also taken into consideration. For calculating the transistor OFF state current, the equation given in [8] will be used:

$$I_{OFF} = \hat{I} e^{K_1 * V_G + K_2 * V_D - K_3 * V_S}, \quad (3)$$

where $\hat{I}$ is the leakage current under normal terminal voltages, K_1, K_2 and K_3 are technology-dependant constants.

When the "Save" signal goes to high value the transmission gate (TG) in Fig. 1 opens, sinking the current from master to retention latch increasing the potential of point B and decreasing the potential of point A. Until the potential of point B has not reached the switching point of two inverters, the feedback does not allow the latch to switch. During this period the master latch current increases the potential of point B until its potential exceeds the switching point, after this the feedback rushes the inverters to switch. For the potential of point B to rise the switching point, the parasitic capacitance of the retention latch should be charged. These parasitic capacitances are represented as a single capacitance placed at the output of TG, Fig.1. That capacitance can be found by summing the input and output capacitances of the inverters in latch. Using the inverters input, output capacitance equations given in [9], the total capacitance will be:

$$C = \frac{5}{2}(C_{OXN}(L_N^M * W_N^M + L_N^R * W_N^R) + C_{OXP}(L_P^M * W_P^M + L_P^R * W_P^R)), \quad (4)$$

where M refers to master and R to retention latch parameters. The charging current of that capacitance can be found by summing all the currents flowing through the ON and OFF transistors, these currents are shown in Fig. 1. Overall the charging current of the capacitor is equal to

$$\begin{aligned} I_{Supp}^M &= I_{P,trd}^M - I_{N,off}^M, \\ I_{Wst}^R &= I_{N,trd}^R - I_{P,off}^R, \\ I_{Chrg} &= I_{Supp}^M - I_{Wst}^R, \end{aligned} \quad (5)$$

where I_{Supp}^M is the current which supplies the master latch, I_{Wst}^R is the current which is spending the retention latch and I_{Chrg} is the parasitic capacitance charging current. Charging current is calculated by subtracting from master latch supply current the retention latch spending current. Using (2), (3) the charging current can be found, which is:

$$\begin{aligned} I_{Chrg} = \frac{1}{2}(&\frac{\mu_P C_{OX,P} W_P^M (V_{DD} - \phi_D^M)(V_{DD} - 2*V_{TH,P} + \phi_D^M)}{L_P^M} + \mu_N C_{OX,N} W_N^R \phi_D^R * \\ &* \frac{(2*V_{TH,N} - 2V_{DD} + \phi_D^R)}{L_N^R} + \frac{2*\exp\{K_1 V_{DD} + K_2 \phi_D^R\} \hat{I} W_P^R}{L_P^R} - \frac{2*\exp\{K_1 \phi_D^M\} \hat{I} W_N^M}{L_N^M}). \end{aligned} \quad (6)$$

To calculate the time of capacitor charging, an assumption should be made that charging current of the capacitor remains constant during the flop data transferring process. Such assumption is justified by the fact that it makes very little divergence in the final result and simplifies the save time equation. Based on this the capacitor charging time can be written as product of capacitance and voltage divided by current.

Using the capacitor (4) and charge current equations the following can be written for retention data saving time of RF:

$$\begin{aligned} t = &\frac{5V_{DD} L_N^M L_P^M L_N^R L_P^R * (C_{OX,N}(L_N^M W_N^M + L_N^R W_N^R) + C_{OX,P}(L_P^M W_P^M + L_P^R W_P^R))}{(\mu_P C_{OX,P} L_N^M L_N^R L_P^R W_P^M (V_{DD} - \phi_D^M)(V_{DD} - 2V_{TH,P} + \phi_D^M) + L_P^M (\mu_N C_{OX,N} * \\ &* L_N^M L_P^R W_N^R \phi_D^R * (2 * V_{TH,N} - 2V_{DD} + \phi_D^R) - 2\hat{I} L_N^R (\exp\{K_1 \phi_D^M\} L_P^R W_N^M - \\ &- \exp\{K_1 V_{DD} + K_2 \phi_D^R\} W_P^R L_N^M))), \end{aligned} \quad (7)$$

where M refers to master and R to retention latch parameters. The ϕ_D^R is the potential of point B and ϕ_D^M is the potential of point A. Based on (7) the saving time of RF can be estimated. After inserting the values of transistor oxide capacitances, which are known

from foundry provided technical specs, inserting the value of supply and threshold voltages and the initial potential values of A and B points, the following for expression (7) we will have the following simple expression:

$$t \approx L^2\left(a + \frac{1}{\frac{k * W_N^R}{W_N^M} - b}\right), \tag{8}$$

where a, k and b are constant numbers depending on specific technology. If the technology is known then these constants can be calculated like it is done in [3] reference. Various dependencies of saving time from master and retention latch parameters can be derived by using expression (7).

Similarly the same way expression for data restoring time can be calculated. If in Fig. 1 the master latch is assumed as retention and retention latch is assumed as master latch then the (7), by changing all M suffixes to R and all R suffixes to M, will turn the expression for data restoring time.

Boosting mechanism for retention and restoration signal generation. In [6] new self retention and restoration mechanism for RF was presented. The idea of that mechanism is adding some voltage sensitive circuit to the flop architecture. That sensor logic can sense the supply voltage changes of the flop and based on that changes identify the power down and power active states. When one of such states is identified the sensor will automatically perform retention data saving and restoring. Such mechanism will eliminate the need for having two save and restore control pins in flop, which will bring to decreasing the flop wiring and as a result it will drastically decrease the chip power consumption.

However the proposed mechanism has one drawback, which is when the power shuts down the voltage value of virtual ground rises towards the supply voltage V_{DD}, and in general that rise is not very quick. As a result of this, the sensor spends lots of time for catching the power down and on state which decreases RF speed. It was estimated, based on (7) that proposed self retention and restoration mechanism can worsen the flop save and restore times nearly by 13.7%. Such a number comes from the fact that virtual ground potential rises towards the supply voltage very slowly which brings to the increase of RF save and restore time.

To overcome this issue, a large PMOS transistor is placed between the real Vdd and virtual Vssv power. In footer gated state this PMOS transistor is operated for a very short period raising the voltage of Vssv virtual ground. This accelerate the retention signal generation. Fig. 2 shows how the PMOS transistor is placed in the design:

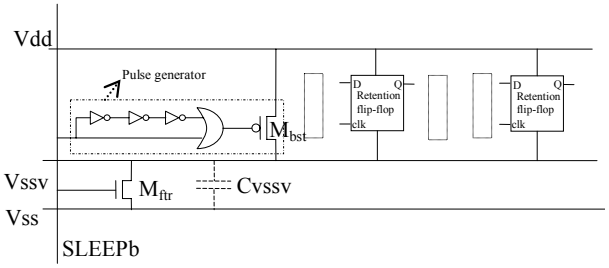

Fig. 2. Boosting transistor placement in design

When the footer gating is performed the SLEEP signal goes to high value, which brings its inverted value SLEEPb to low value. This causes closing of the M_{ftr} footer transistor and performing power gating. When SLEEPb signal goes to low value the "Pulse generator" circuit generates a short pulse which opens the M_{bst} boosting PMOS transistor for a short time. That boosting transistor passes some charge to virtual ground increasing its potential. This accelerates the RF sleep signal generation.

The size of PMOS transistor and the pulse width are important design considerations, since they determine the amount of time the virtual ground takes to reach its steady state and the amount of energy dissipated during that period. It can be shown that the total amount of charge supplied by pulse generator to virtual ground (V_{ssv}), denoted by Q_{scp}, is proportional to the PMOS transistor saturation current and the pulse generator pulse width.

$$Q_{pg} \infty I_{d,sat} P_w , \tag{9}$$

where Q_{pg} is the amount of supplied charge, $I_{d,sat}$ is the PMOS transistor saturation current, P_w is the pulse width. Using the saturation current is given in [7]:

$$I_{d,sat} = \frac{1}{2}\mu_p C_{ox} \frac{W}{L} (V_{dd} - V_t)^2 , \tag{10}$$

where μ_p is the carrier mobility, C_{ox} is the oxide capacitance, W and L are the channel width and length of pulse generator PMOS device, V_t is the threshold voltage. From (9) and (10) for total amount of charge the following expression can be written:

$$Q_{pg} = mWP_w , \tag{11}$$

where m is proportionality constant. The value of m can be obtained by simulating a particular size of PMOS device and monitoring the amount of charge supplied by the device while varying the pulse width. By solving (10) the total amount of charge that needs

to charge the V_{ssv} to V_{dd} can be found.

Let the capacitance involved in V_{ssv} be denoted by C_{vssv}. This capacitance is shown in Fig. 2. The amount of charge stored in the capacitance when the footer transistor M_{ftr} is off and M_{bst} is on, is denoted by Q_{vssv}, hence using the basic charge equation stored in capacitor the following can be written:

$$Q_{vssv} = C_{vssv} V_{dd} . \tag{12}$$

Obtaining the Q_{pg} from (11) and Q_{vssv} from (12) the two are equated, which yields the product of pulse generator size (W) and pulse width (P_w)

$$WP_w = \frac{C_{vssv} V_{dd}}{m} . \tag{13}$$

Any combination of W and P_w that satisfies (13) will serve the purpose. The charge which rises the potential of virtual ground towards the supply voltage is directly proportional to the current which flows from supply to virtual ground and the amount of time during which the current is flowing. This basic principle is shown in (13). Transistor current is directly proportional to transistor width (W) and the time of current flowing is equal to pulse width (P_w), the designer needs to have those two values to meet the design targets. However, if the pulse width P_w has too small value, then too many pulse generators need to be added in the scheme, besides, the pulse should be wide enough to avoid distortion of pulse shape over long distance.

Experimental results. In the first part of the experiment the equations for retention data saving time and retention logic power consumption were analyzed. The accurate HSPICE [10] simulations were performed for various types of RF, and save time values for that flop were calculated. Then the save time value was calculated for the same flop, using the equation presented in (7). Fig. 3 shows the save time expression for different values of master and latch transistor width relations and comparison with HSPISE results.

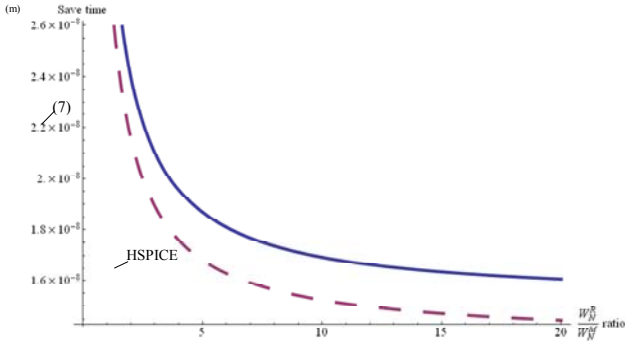

Fig. 3. Save time comparison results

The dashed plot corresponds to HSPICE simulations and the thick one is the one that was received by using equation (7). It was estimated that the analytical results differ from the HSPICE simulation results not much than 14.7%. This is the acceptable value, as the main purpose is to have some estimation of save time dependency from flop parameters.

Using the (7) expression, various dependencies of RF save time parameter from flop transistor parameters and technology parameters were obtained.

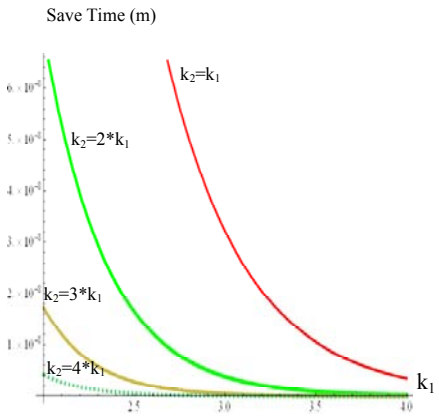

Fig. 4. Save time dependency from K technology parameters

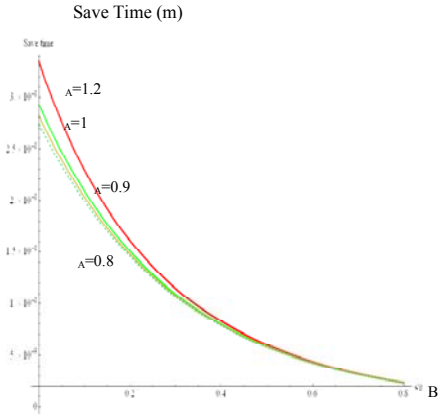

Fig. 5. Save time dependency from potentials of A and B points

In Fig.4 the save time dependency from the variation of k_1 parameter is shown for different values of k_2.Mainly the k_1, k_2 parameters reflect the transistor threshold change depending on productions. So the results in Fig. 4 show that if during the production the threshold voltages of transistors increased, then this will bring to save time decrease. Fig. 5 shows save time dependency from the potential of B point in Fig.1 for different values of A potential. By increasing the initial threshold potential of B point, the save time is decreased because it is easier for master latch to write the data in retention logic.

The next part of experiments is connected with the proposed boosting mechanism. The virtual ground voltage change was observed for different PMOS pulse width values. Figure shows the virtual ground, V_{ssv}, voltage rise comparison between the scheme which uses the boosting mechanism and the scheme which does not use it.

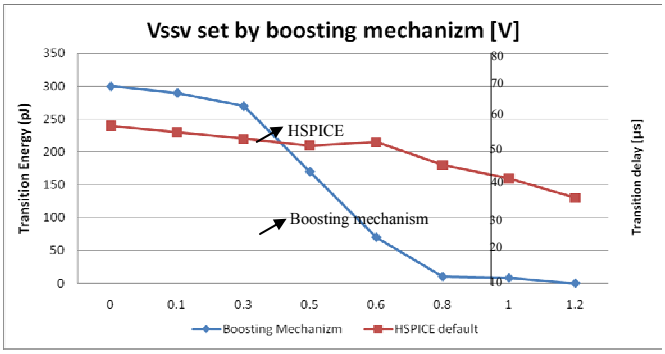

Fig. 6. Boosting mechanism in comparison with default method

In this plot the proposed method is applied on the memory controller. The boosting mechanism pulse width and boosting transistor size is varied during the experiment. In Fig. 6 the x-axis corresponds to the potential of V_{ssv}, the y-axis on the right-hand side indicates the transition delay, which is the interval from turning off the footer to the point at which V_{ssv} settles down to its steady-state potential, which is assumed to fall within 5% of supply voltage. The y-axis on the left-hand side indicates the total energy dissipated during the transition delay. As it can be seen from the plot, the proposed boosting mechanism increases the speed of virtual ground voltage rise.

Also, the proposed mechanism was tested on various benchmark circuits, the pulse width and transistor sizes were estimated. All the benchmarks were taken from ISCAS family [11]. Below the table shows the results of measurements:

Table

Measurement results with usage of boosting mechanism in the design.

Name	#Gates	#FFs	#POs	#PGs	#Bst Mchs	#Footers	P_w (ps)	W (ps)	V_{ssv} set by Bst mech.
s35932	3513	1728	320	14	53	37	751	432	1,15
s38417	3333	1564	106	13	50	28	817	374	1,139
s38584	4294	1275	304	13	52	33	902	353	1,148
b12	855	119	6	2	6	4	729	50	1,157
aes1	503	228	129	2	8	7	759	65	1,149

Based on this Fig. it can be seen that flop transition time increases rapidly which brings to drastic increase of flop self save and restore operation speed. Overall this speed increase brings to power saving, as the switching process in flop takes less time. As a result of increase in self save and restore operation, the flop switching time is decreased which brings in average 32.6% power saving.

Conclusion. The derived expressions differ from HSPICE results not more than 14.7%. Based on those experiments various dependencies of RF save, restore time from flop other parameters, e.g. RF transistor sizes, technological parameters, flop nodes potential values in sleep mode were presented. Boosting mechanism for RF self saving and restoring capability helps significantly increase the self save and restore time of RF, due to increase in virtual ground voltage rising speed. Such speed increase brings to power saving in the whole design. The experiments show that with the proposed mechanism nearly 32.6% power saving can be reached. The drawback of the boosting mechanism is thr increase of design time and verification complexity.

REFERENCES

1. **Roy K., Mukhopadhyay S., Mahmoodi S.** Leakage current mechanisms and leakage reduction techniques in deepsubmicrometer CMOS circuits // IEEE transactions on Very Large Scale Integration (VLSI) Systems.- Feb. 2003.-Vol. 91, no. 2.-P.305-327.
2. **Swarup B., Saibal M.** Low-Power Variation-Tolerant Design in Nanometer Silicon.- Springer, Dordrecht, 2011.
3. **Leah S., Jason B.** Managing Power in 45nm and 65nm Designs // 9th International Synopsys Users Group (SNUG) Conference.-2009.-P. 730-733.
4. **Bing B., Augusli K., Kun W.** Fault modeling and testing of retention flip-flops in low power designs // Proceedings of the 2009 Asia and South Pacific Design Automation Conference.- 2009.- P. 684-689.
5. **Kim H., Shin Y.** Semicustom design methodology of power gated circuits for low leakage applications // IEEE transactions on computer-aided design of integrated circuits and systems.-2009.-Vol. 54.-P. 512-516.
6. **Melikyan V., Petrosyan H., Durgaryan A., Dragan T.** New Retention Flop Architecture with Phase Frequency Detection Capabilities // 55th Conference for electronics, telecommunications, computers, automatic control and nuclear engineering (ETRAN).-Serbia, 2011.-P. 58-62.
7. **Behzad R.** Design of Analog CMOS Integrated Ciruits.-McGraw Hill, Boston, 2001.
8. **Hao X., Ranga V., Jone W.** Dynamic Characteristics of Power Gating during Mode Transition // IEEE transactions on Very Large Scale Integration (VLSI) Systems.-Feb. 2011.-Vol. 19.-P. 237-249.
9. **Backer J.** CMOS Circuit Design, Layout and Simulation // Willey IEEE Press Series on Microelectronic Systems.-1997.
10. http://www.synopsys.com/Tools/Verification/AMSVerification/CircuitSimulation/HSPICE
11. http://ieee-cas.org/

SEUA (POLYTECHNIK). The material is received 20.09.2011.

Վ.Շ. ՄԵԼԻՔՅԱՆ, Հ.Պ. ՊԵՏՐՈՍՅԱՆ, Ա.Ա. ԴՈՒՐԳԱՐՅԱՆ, Է.Հ. ԲԱԲԱՅԱՆ,
Ն.Խ. ԱՍԼԱՆՅԱՆ

ՎԻՃԱԿԻ ՊԱՀՊԱՆՄԱՄԲ ՌԵԳԻՍՏՈՐՆԵՐԻ ԻՆՔՆԱՊԱՀՊԱՆՄԱՆ/ՎԵՐԱԿԱՆԳՄԱՆ ԼԱՎԱՐԿՄԱՆ ՄԵԽԱՆԻԶՄ

Սպառվող ստատիկ հզորության նվազարկման ամենատարածված եղանակը սնման լարման շրջափակումն է: Այդպիսի շրջափակման ընթացքում սխեմայի ընթացիկ վիճակը հիշելու նպատակով օգտագործվում են վիճակի պահպանման ռեգիստորներ (ՎՊՌ): Դրանք ղեկավարվում են արտաքին սնուցման ղեկավարման հանգույցով: Արդյունքում՝ սխեմայում մեծանում են միջմիացումների երկարությունները և բուֆերների քանակը: ՎՊՌ-ների բնութագրական պարամետրերի գնահատման համար ստացված են բանաձևեր: Հետազոտված են այդ պարամետրերի կախվածությունները ՎՊՌ-ի տրանզիստորների չափերից, տեխնոլոգիական գործոններից, ՎՊՌ-ի հանգույցների պոտենցիալների արժեքներից: Առաջարկված է ՎՊՌ -ների ինքնապահպանման և վերականգման հատկության լավարկման նոր եղանակ: Այն ապահովում է հանգստի վիճակին անցման ընթացքում ՎՊՌ -ի զերծպահումը կարճ միացման հոսանքից: Միաժամանակ, մեկ կարգով մեծացնում է ռեգիստորի ինքնապահպանման և վերականգման ժամանակը: 45 *նմ* տեխնոլոգիայով իրագործված ISCAS շարքի տարբեր սխեմաների փորձարկումների արդյունքները ցույց են տվել, որ առաջարկված եղանակը փոքրացնում է ՎՊՌ-ի ինքնապահպանման և վերականգման ժամանակը, որը իր հերթին հանգեցնում է 32,6% հզորության խնայման:

Առանցքային բառեր. սնուցման լարման շրջափակում, վիճակի պահպանման ռեգիստոր, ստատիկ հզորություն, պահման ժամանակ, վերականգնման ժամանակ, կորուստային հոսանք:

**В.Ш. МЕЛИКЯН, Г. П. ПЕТРОСЯН, А.А. ДУРГАРЯН, Э.Г. БАБАЯН,
Н.Х. АСЛАНЯН**

МЕХАНИЗМ УЛУЧШЕНИЯ САМОСОХРАНЕНИЯ/ВОССТАНОВЛЕНИЯ УДЕРЖИВАЮЩИХ РЕГИСТРОВ

Наиболее распространенным способом уменьшения потребляемой статической мощности является блокировка напряжения питания. С целью запоминания текущего состояния схемы в течение такой блокировки используются регистры удержания состояния (РУС). Эти регистры управляются внешним узлом управления питанием. В результате увеличиваются длины межсоединений и количество буферов в схеме.

Получены аналитические выражения определения характерных параметров РУС. Исследованы зависимости этих параметров от размеров транзисторов РУС, технологических факторов, значений потенциалов узлов РУС. Предложен новый способ улучшения свойств самосохранения/восстановления РУС. Способ обеспечивает устранение РУС от тока короткого замыкания во время перехода в режим ожидания. Одновременно на один порядок увеличивается время самосохранения/восстановления. Результаты испытания различных схем из ряда ISCAS, изготовленных по 45*нм*-овой технологии, показывают, что предложенный способ увеличивает время самосохранения/восстановления, и, как следствие, снижается энергопотребление на 32,6%.

Ключевые слова: блокировка цепи питания, удерживающие триггеры, статическая мощность, динамическое энергопотребление, время самосохранения/восстановления, ток утечки.