A Survey on Leakage Power Reduction Techniques by Using Power Gating Methodology

Pramod Kumar. M.P #1, A.S. Augustine Fletcher #2

#1 PG scholar, VLSI Design, Karunya University, Tamil Nadu, India
#2 Assistant professor, ECE Department, Karunya University, Tamil Nadu, India

Abstract—As the technology moves into deep sub-micron region, the power consumption of the integrated circuit will be more. In the current technologies, the leakage power is the major part in the total power consumption. Power gating is a technique which is used to reduce the leakage power by shutting off the idle logic blocks using sleep transistors. Different power gating methods are available now. These help in reducing the power, delay and switching time of the logics. This survey paper mentions some important power gating techniques and its comparison.

Keywords—CMOS, MTCMOS, power gating, threshold voltage, sleep mode etc.

I. INTRODUCTION

According to Moore’s law, the number of transistors including in a chip be doubled in every 18 months. So if the technology is down scaling the area of each device in a chip reduces. Lesser area increases the power consumption. There are lot of methods available to reduce the power in a chip. Total power consumption will be divided into static and dynamic. Static power is the power loss at the time of transistors are working in weak inversion region. Dynamic power relates to the charging and discharging behaviour of capacitors.

Power gating is a technique which is used to reduce the power when the logics are in idle state. Power gating based on the Virtual VDD and Virtual ground concepts. An NMOS sleep transistor is connected in between the circuit ground and actual ground is named as virtual ground. These NMOS sleep transistors are named as footer transistors. Virtual VDD concept is one in which a PMOS transistor is inserted in between the actual VDD and circuit VDD. Controlling the logic circuit by using an NMOS sleep in power gating techniques. The footer transistor is switching ON and OFF because of its applied gate voltage. The important parameter is the sizing of sleep transistors.

Some important power gating techniques are single mode power gating, tri-mode power gating, multiple threshold CMOS technique, multiple sleep mode power gating, hybrid multiple mode power gating etc. In this paper, we are mentioning some other power gating techniques also. In single mode power gating, we are inserting an NMOS transistor in between the circuit ground and actual ground. We are applying different voltages into the gate of this footer transistor. If the applied gate voltage is above the threshold voltage of the footer, the transistor is ON and it works as a low resistance path. If the gate voltage is less than the threshold voltage of the footer, it works in weak inversion region and it provides as a high resistance path to ground.

In the tri-mode technique, we are adding a PMOS transistor in parallel to the NMOS sleep. The number of modes are increased by using this technique. The footer works in three different modes and it helps to reduce the wake up power and power penalty. Large amount of static power is also reduced. MTCMOS is the multiple threshold CMOS power gating technique and a PMOS header and an NMOS footer is added into the logic. In this technique, low threshold voltage (VTH) transistors are used in logic cell and high VTH transistors are used as header and footer transistors.

Multiple sleep power gating is one which the footer transistor is biased with different gate voltages. These different gate voltages which are less than threshold voltage of the footer and the footer remains staying in weak inversion region. Here, we are using four different modes and it helps to reduce the transition time and power also. Hybrid multiple mode power gating increases the sleep modes to eight and it takes very less power compared to the other existing techniques.

Considering the size of the sleep transistor, it may modelled as a resistor RACTIVE [1]. The generated voltage is equal to IACTIVE * RACTIVE. Where IACTIVE is the total current of the logic which works in active region. The virtual ground voltage is termed as VGND. The relation between threshold voltage and propagation is

\[ T_{PD} \alpha C_L \cdot V_{DD} / (V_{DD} - V_X - V_{TH}) \alpha \] (1)

Where \( C_L \) is the load capacitance, \( V_X \) is the voltage drop across the sleep transistor, \( V_{TH} \) is the threshold voltage, \( T_{PD} \) is the propagation delay and \( \alpha \) is the modelling constant.

The relationship between VGND [2] and WSLEEP can be written as

\[ W_{SLEEP} = L / (\mu N \cdot C_{OX} \cdot R_{ON} \cdot (V_{DD} - V_{TH})) \] (2)

Where L is the channel length and \( V_{TH} \) is the threshold voltage, \( \mu N \) represents the mobility of the electrons and \( C_{OX} \) represents the cell capacitance per unit area of the sleep transistor. Finally, \( R_{ON} \) denotes the channel resistance of the sleep transistor operating in the linear region.

Singh et al. derived the wake up time [3], as follows:
The wake-up time is given by:

\[ T_{\text{wake-up}} = C_{\text{CIRCUIT}} \cdot \frac{V_{\text{GND}}}{I_{\text{ON,F}}} \]  

(3)

Where \( C_{\text{CIRCUIT}} \) is the circuit capacitance and \( I_{\text{ON,F}} \) is the on current of the footer.

II. A MULTI-MODE POWER GATING STRUCTURE FOR LOW VOLTAGE DEEP-SUBMICRON CMOS ICs [4]

This paper proposes a power gating structure that supports both an intermediate power saving and data retaining mode. Test structures are fabricated in 130nm CMOS bulk technology. Tri-mode power gating allows a choice between a large reduction in leakage without state retention and an intermediate level of leakage reduction with state retention. A reduction in the ground bounce noise is induced by power mode transitions of the power gating structure. The leakage current is reduced when the ground supply to the logic circuit is interrupted by the small NMOS sleep transistor and is moderately reduced when the PMOS data retention switch is used to reduce the rail to rail voltage. This power gating structure were to be applied to a SRAM, required to retain data despite process, PVT variations and supply voltage [5], [6].

In this technique, adding a single PMOS in parallel to a power gating structure with NMOS transistor, which leads to a power gating structure that can support an additional intermediate power saving and data retaining mode. Fig. 1 shows tri-mode power gating structure with RUN/IDLE mode.

Here PG is asserted high to force the NMOS transistor in the power gating structure into a low resistance state, while HLD is set high. NMOS is used to short the virtual ground \( V_{\text{GND}} \) of the logic circuit to the real ground potential, allowing the full supply voltage to be applied across the circuit. This permits a high speed operation. In COLD mode, the state does not retain. PG is held low while HLD is high, current path to GND is blocked. The voltage across the logic circuit collapses, suppressing both gate and sub threshold leakage currents. In the state retention or PARK mode, both PG and HLD are asserted low. The NMOS device is turned off and the PMOS operates as a source follower. The \( V_{\text{GND}} \) is held at a voltage \( V_{TP} \) above that of the ground rail, where is the threshold voltage of the PMOS in the PARK mode. In this mode, the \( V_{\text{GND}} \) voltage level is limited by \( V_{TP} \). As a result, the state is retained and the ground bounce induced by power mode transitions is smaller than it is in COLD mode. PARK mode reduces the ground bounce induced by the transition from COLD to RUN/IDLE.

The reduction in leakage power is shown to be 2.6 times less than in IDLE mode. Compared results show that the power mode transition from COLD to IDLE through PARK reduces the ground bounce by up to 50%, depending on the supply voltage.

III. DYNAMIC AND LEAKAGE POWER REDUCTION IN MTCMOS CIRCUITS USING AN AUTOMATED EFFICIENT GATE CLUSTERING TECHNIQUE [7]

This paper presents efficient gate clustering methods in MTCMOS circuits by modelling the problem via Bin-Packing and Set-Partitioning. Circuit’s routing complexity into consideration in Set-Partitioning technique which is critical for deep sub-micron technology. Performance is achieved sufficiently, while reducing the overall sleep transistors area. The main objective of the pre-processing stage is to group gates into sub clusters such that the combination would not exceed the maximum current of any gate within the cluster. Heuristic forms a set of sub clusters of gates that when combined would not exceed the maximum current of any gate within the cluster.

The Bin-Packing (BP) technique is very efficient when it is applied to small circuits that have unbalanced structures. Limitation is that the BP technique does not take the physical locations of the gates on the chip into consideration. This might cause two gates located far apart to be clustered together which will augment the routing complexity for larger circuits. The Set-Partitioning (SP) technique solves this and consequently reduces the routing complexity of the circuit [8].

BP pre-processing algorithm has a worst case complexity \( O(n^2) \), where ‘n’ is the number of gates in the circuit. But SP algorithm complexity is \( O(nk) \), where ‘n’ is the number of gates in the circuit and ‘k’ is the maximum gates to be appended in a cluster. It is recommended that heuristic search techniques such as Genetic Algorithms would be used instead of the CPLEX solver for large circuits.

In this paper, six benchmarks are used as test vehicles i.e. a 4-bit Carry Look-Ahead adder, a 32-bit parity checker, a 6-bit array multiplier design, a 32-bit Single Error Correcting circuit (C499 ISCAS-85 benchmark), a 4-bit ALU/Function Generator (74181 ISCAS-85 benchmark) and finally a 27-bit Channel Interrupt Controller (C432 ISCAS-85 benchmark).

SP technique achieves 88% and 66% leakage reduction compared to [8] and [9]. The location of the blocks in order to reduce the overall interconnects, providing more optimization to the area are the main advantages of the SP technique. The proposed technique achieves on an average of 15% savings for dynamic power and 90% savings for leakage power. On average, the BP technique reduces dynamic by 15% and leakage power by 90%.
IV. A ROBUST POWER GATING STRUCTURE AND POWER MODE TRANSITION STRATEGY FOR MTCMOS DESIGN [10]

This paper introduces an approach for reduction of transition time from sleep mode to active mode while assuring power integrity for the rest of the system. The problem during the sleep to active mode transition is to minimize the wakeup time while constraining the current flowing to ground. During the process, also another important objective is considered i.e. limiting the number of sleep transistors. This technique is to achieve the mentioned objectives based on effectively clustering logic cells and scheduling wakeup signals for the clusters. The algorithms provided in this paper have low computational complexity and very effective.

The size of sleep transistors can be determined by using well known methods [7], [11]. In the Wakeup Signal Scheduling (WSS) Problem, clustered the logic cells into a minimum number of clusters and found the optimum turn-on times for logic clusters in the circuit. So minimized the overall turn-on time of the circuit. Solved the wakeup signal scheduling problem by solving each of the clustering and scheduling the problems separately and sequentially. The heuristic used for clustering returns only one solution out of many possible solutions. The result of scheduling depending on the solution provided by the clustering algorithm in this section proposed a new technique where clustering and scheduling are done simultaneously and the overall objective is targeted continuously throughout the algorithm. This technique is also a heuristic but tends to produce better results compared to the two step approach.

This is used to find the delay and current profile of each logic cell in a 180nm standard cell library for a given sleep vector for all possible input combinations to the logic cell. Also applied that algorithm to a number of circuits from the ISCAS test benchmark suite. Observed that the ground current dominates the supply current for techniques that only use NMOS sleep transistor.

V. ENHANCED MULTI-THRESHOLD (MTCMOS) CIRCUITS USING VARIABLE WELL-BIAS [12]

Reverse bias on the interrupt switch during sleep mode can allow substantial leakage reduction without the expense of adding a high threshold device. This switch is constructed with a low \( V_{TH} \) device. Adding a forward bias to the power supply interrupt switch during active mode of operation will reduce the performance penalty associated with MTCMOS. Increasing performance in MTCMOS is boosted by the gate drive.

In MTCMOS logic, the circuit is gated by a high threshold header and footer switch and the logic is constructed by low threshold transistors. Two layout styles were examined for inserting the power supply interrupt devices i.e. an integrated method and an external ring. In the integrated method, which places a footer within each standard cell row at the intersection of the M1 virtual ground and the M2 real ground. Sizing the footer by parallel placement of footer books can be based on the power density limits. Since the footer is placed in each row alongside the standard cells, a common p-well is used for all NFETS within each row. The same methodology could be extended to PFET header switches, where stripes of n-wells isolate each row of PFETS in the gates.

An alternate approach is available using an external ring for inserting power supply interrupt devices. Here, the power distribution grid is interrupted at the M2 and M3 by two rings, the inner ring supplying virtual ground and the outer ring connecting to the real ground. Area between the rings can be used for the footer switch. The external ring approach has the benefits of easily allowing the power interrupt switch to be placed in a separate well from the internal logic, enabling separate body connections. This will eliminate increased reverse body bias on the switched logic NFETs, complicating p-well contact routing and routing real GND to internal non-switched logic. Another benefit of the external ring approach is that pre-existing fixed layout cores can more easily be converted to a power interrupt switch.

A low \( V_{TH} \) PFET header can allow better performance and lower active power over a high \( V_{TH} \) header device. By applying a large reverse body bias to the header during standby state, similar leakage characteristics can be achieved. The header with high \( V_{TH} \) device shows that the leakage can be reduced by almost 100x by reverse body bias, being limited by GIDL [13] effects resulting from large reverse bias voltage.

VI. LEAKAGE POWER REDUCTION THROUGH HYBRID MULTI-THRESHOLD CMOS STACK TECHNIQUE IN POWER GATING SWITCH [14]

In this paper, two hybrid digital circuit design techniques are proposed, hybrid multi-threshold CMOS complete stack technique and hybrid multi-threshold CMOS partial stack technique for reducing the leakage power dissipation in mode transition. This paper combines the advantages of both stack and MTCMOS techniques. Tri-modal switch’s performance depending on these techniques reduce the leakage power. These techniques are implemented using Cadence virtuoso tool to find the leakage power dissipation and propagation delay. Thus, the stack techniques are used in the tri-modal MTCMOS switch design enabling active, drowsy and sleep modes. Leakage power reduction in the mode transition is reduced. The propagation delay of this technique is higher than the MTCMOS technique.

The drowsy mode reduces the leakage current while preserving the content of the cell [15]. Reduced the leakage power during mode transition in the various modes at the same circuit. The tri-modal switch operates in the sleep or drowsy mode depending on the value of the DROWSY signal. These hybrid techniques provide an improved performance in terms of leakage power compared with the other techniques.

In the hybrid multi-threshold CMOS complete stack technique, a high \( V_{TH} \) PMOS sleep transistor is inserted between \( V_{DD} \) and the pull up network and a high \( V_{TH} \) NMOS transistor is inserted between the pull down network and ground. Stacking of all transistors are done by replacing transistor width \( W \) with two series connected transistors of width \( W/2 \). In the hybrid multi-threshold CMOS partial stack technique, the high threshold header and footer transistors are
stacked. Stacking of low $V_{TH}$ NMOS and low $V_{TH}$ PMOS transistors of the logic circuit is not performed. Only partial stacking is done to reduce the circuit propagation delay in active mode. Fig 2 shows the logic circuit diagram of hybrid MTCMOS complete stack technique.

![Fig 2. Hybrid MTCMOS complete stack technique [14]](image)

These hybrid technique provides a reasonably low leakage solution than others. Hybrid MTCMOS complete stack technique has the higher propagation delay, but it is the most appropriate choice than others. The delay calculations are done by using Cadence tool. Hybrid MTCMOS partial stack delay is lower than the complete stack technique.

VII. ROW BASED POWER GATING:- A NOVEL SLEEP TRANSISTOR INSERTION METHODOLOGY FOR LEAKAGE POWER OPTIMIZATION IN NANOMETRE CMOS CIRCUITS [9]

Sathanur et al. proposes a layout aware methodology that facilitates sleep transistor insertion and virtual ground routing on row based layouts. Also introduces a clustering algorithm that is able to handle timing and area constraints, and extended it to the case of multi threshold sleep transistors to increase leakage savings. This row based power gating methodology favours fast design closure and makes it suitable for the implementation as a CAD tool. Finally developed a multi threshold sleep transistor synthesis technique which enables to further reduce the total leakage than other sleep transistor techniques.

The contributions can be summarized below:

1. New algorithms for optimally determining a sub set of the gates in the design to be power-gated taking into account both the timing and the area overhead.
2. A post-layout sleep transistor insertion methodology, which assumes that all sleep transistors are placed in dedicated rows. Thus, the leakage reduction is easily estimated and the insertion approach incurs minimal layout modifications.
3. This row-based sleep transistor insertion methodology to be used as a design exploration tool.
4. Multi-threshold sleep transistor synthesis technique, which enables to design an optimal sleep transistor with the available threshold voltages in the library with improved leakage savings.

This partial power gating scheme have many similarities with other timing-driven power optimization techniques, such as multi-$V_{DD}$ and multi-$V_{TH}$ design [16]. The cells that are not standing on the critical path are replaced with slower, but more dynamic and static power efficient ones while meeting the original input constraints. The gating of a row results in an increase in delay of all the cells in that row. This amount of delay penalty is affected by the area constraint.

Timing and area constrained clustering is one of the key steps in this partial power gating methodology. It consists of an iterative process. The best clusters of layout rows that will lead to maximum leakage savings under such constraints. The virtual ground voltage depends on the peak discharge current of the cluster. Fig. 3. Shows the standard cell layout after clustered power gating.

![Fig. 3. Standard cell layout after clustered power-gating [9]](image)

This row based partial power gating methodology to some of the largest benchmarks taken from the ISCAS 85 and ISCAS 89 suites, as well as to some industrial designs. Each circuit was synthesized and placed using a 65nm CMOS technology using Synopsys Physical Compiler for optimal timing.

VIII. POWER GATING WITH MULTIPLE SLEEP MDES[17]

This paper proposes four different sleep modes and less transition time between ON and OFF. So, the leakage power is gradually reduced. An NMOS footer transistor is placed below the logic and it is driven by four different gate voltages. A robust bias generator circuit is created and two voltages are taken from it, which is less than the threshold voltage of the footer transistor. The bias generator is shown in Fig. 4.

Bias generator circuit consists of PMOS and NMOS transistors. Transistor M1, M2 and M3 are equally sized and form current mirror circuit which is biased by VGS of M1. Transistors M4 –M7 are equally sized and width of M8 is almost four times than M7. The biasing voltage of the power gating circuitry, named $V_{BIAS}$ is taken across M8. The size of M8 is again changed and taken one more $V_{BIAS}$ voltage. The
intermediate gate voltages (V1, V2, V1<V2<V_{TH}) for Sleep and Dream modes are generated using bias generator circuit.

Fig 4. Robust Bias Generator circuit [17]

A 4:2 decoder is used to select the different voltages. These voltages are applied into D0-D3 as shown in Fig 5 and any one of the transistor is ON at a time. D1 and D2 acts as pass transistor logic. The footer mainly works in weak inversion region. The wake up power penalty in each transition is reduced. So, overall power consumption is less by the use of this power gating technique.

Fig 5 Logic diagram of multiple mode power gating [17]

Tested the robustness of above circuit against process-voltage-temperature variations. Generated seven PVT corners by selecting various combinations of V_{DD} (0.9, 1.0 and 1.1V), temperature (55C, 85C and 115C) and process (weak, nominal and best) values. As expected, the absolute values of leakage and wake up overhead showed a significant change with PVT variations. The proposed circuit is robust, if it ensures that the relative trade-off between leakage and wake up overhead in various sleep modes.

IX. A NOVEL HYBRID MULTIPLE MODE POWER GATING [18]

This paper proposes a hybrid power gating technique with multiple sleep modes, each mode represents the trade-off between the wakeup overhead and leakage savings. This hybrid power gating reduces a large static power than existing multimode technique. The design of hybrid power gating technology in 8 bit ALU was done in Cadence Virtuoso tool.

Hybrid multiple mode power gating consists of logic sleep transistors and tri-mode sleep transistors.

In this technique, one NMOS sleep transistor is added below in each logic module. Tri-mode sleep transistors (an NMOS and a PMOS is connected in parallel) and one more logic sleep transistor is connected as in [18]. In this logic, SL1 and SL2 are tri-mode sleep transistors and SL3-SL7 are logic sleep transistors. Static power and leakage power are very much reduced than the existing techniques as well as optimized ALU [19]. Fig 6 shows the static power comparison.

Fig 6 static power comparison [20]

The single bit transition in each mode reduces the ON/OFF time as well as power consumption. Less circuit complexity and area are the other important characteristics. This hybrid multiple mode power gating shows 15% static power reduction than the existing multimode technique in 8 bit ALU. Additional gate count is also reduced.

X. POWER GATING IN 8 BIT ALU: A COMPARATIVE STUDY [20]

This paper presents a comparative study of different power gating techniques. MTCMOS is one which a high V_{TH} header and footer transistors and low V_{TH} logic reduces the power efficiently. Single mode power gating with single sleep transistor and tri-mode power gating with two sleep transistors also reduces the large amount of static power. In tri-mode, 3 different sleep modes and the less ON/OFF transition time helps to reduce the static power. The multiple mode power gating consists of four different modes and hybrid multiple mode consists of 8 different modes.

As the number of modes are increasing, the transition time as well as power are reduced. Hybrid multiple mode is a combination of tri-mode and single mode power gating. Comparison of all this power gating methods in 8 bit ALU and hybrid technique also shows less delay and power delay product. The comparison results are shown in Fig. 7. Hybrid multiple mode saves at least 58% power than other techniques in 8 bit ALU.
XI. Advantages of Power Gating Techniques

Power gating is a very effective approach to minimize standby leakage and propagation delay. It is based on adding the devices called sleep transistors. Virtual VTH and virtual GND concepts are used in power gating techniques. By keeping switching speed high, the leakage power is reduced up to an extent. Effective use of power gating requires proper sizing of the sleep transistors, since it affects the overall performance. Different power gating technologies are available now. MTCMOS technique and single mode power gating reduces the leakage power as well as delay. But MTCMOS faces a difficulty of different threshold voltages. Tri- mode and multiple sleep mode power gating techniques overcome from these issues and reduces the transition time between switching. The row based power gating considers the area also. But its structure is complex. The hybrid multiple mode power gating reduces the static power, delay as well as power delay product.

XII. Conclusion

As the technology is down scaling, the power dissipation in integrated circuit is more. Our aim is to reduce the maximum power without increasing delay and area. In the recent nanometre technologies, the leakage power is more than the dynamic. We are reducing the leakage power with the help of power gating techniques. This survey paper compares the differences in power gating techniques and its advantages.

References