A Semiempirical Model for Wakeup Time Estimation in Power-Gated Logic Clusters

Vivek D. Tovinakere
INRIA/IRISA
University of Rennes 1
Lannion 22300, France
vivektd@irisa.fr

Olivier Sentieys
INRIA/IRISA
University of Rennes 1
Lannion 22300, France
sentieys@irisa.fr

Steven Derrien
INRIA/IRISA
University of Rennes 1
Lannion 22300, France
sderrien@irisa.fr

ABSTRACT

Wakeup time is an important overhead that must be determined for effective power gating, particularly in logic clusters that undergo frequent mode transitions for run-time leakage power reduction. In this paper, a semiempirical model for virtual supply voltage in terms of basic parameters of the power-gated circuit is presented. Hence a closed-form expression for estimation of wakeup time of a power-gated logic cluster is derived. Experimental results of application of the model to ISCAS85 benchmark circuits show that wakeup time may be estimated within an average error of 16.3% across 22× variation in sleep transistor sizes and 13× variation in circuit sizes with significant speedup in computation time compared to SPICE level circuit simulations.

Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids

General Terms

Algorithms, Design, Performance

Keywords

Design automation, leakage current, power gating, wakeup time

1. INTRODUCTION

Power gating has emerged as an important technique to minimize static power and energy consumption in CMOS circuits [3]. As MOSFETs are scaled down to sub-100nm dimensions, an exponential increase in subthreshold leakage current is observed due to the reduction in threshold voltage (V_{th}) to maintain gate overdrive [9]. In ultra-low power circuits with constrained energy budgets, energy consumption due to static currents may dominate its dynamic counterpart for low duty cycle operation. Hence circuit techniques for power gating structures and exploration of power gating opportunities for automated design of power-gated circuits have received significant attention.

A power gating structure cuts-off bias voltages for MOS devices so that bias-dependent leakage current in the logic circuit reduces significantly in standby state. A simple power-gated circuit shown in Fig. 1 and used in this work consists of a high-V_{th} PMOS sleep transistor connected between power supply rail (V_{dd}) and virtual power supply node, Virtual-Vdd (V_{Vdd}) of the logic cluster. A cluster refers to an ensemble of connected logic gates power-gated by a sleep transistor. The gate terminal of the sleep transistor is connected to a control signal $SLEEP$, to switch the sleep transistor between on and off states. A power-gated circuit operates in three modes in a typical power gating cycle as shown in Fig. 2. When $SLEEP$ is high, power supply to the logic is cutoff; V_{Vdd} decreases and the circuit is said to be in sleep mode. The leakage current decreases exponentially with V_{Vdd} resulting in energy savings. When $SLEEP$ is low, current flows through the sleep transistor to charge circuit capacitances. Due to charging effect, V_{Vdd} increases until it reaches a steady state value less than V_{dd}. We refer to this mode of operation as wakeup mode and the mode of operation after wakeup as active mode.

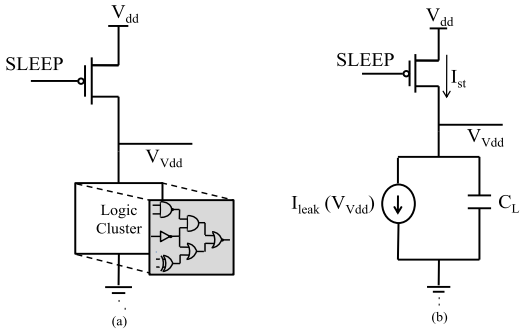

Figure 1: (a) Power-gated logic cluster of header type (b) Equivalent circuit of logic cluster

In this paper, a semiempirical model for V_{Vdd} based on polynomial representation of leakage current in a logic cluster and linear region resistance of sleep transistor is presented. A method to estimate steady-state Virtual-Vdd voltage after wakeup mode using leakage current profiles of constituent logic gates is described. Further, a closed-form expression is derived for estimation of wakeup time of the power-gated circuit. The model for Virtual-Vdd in sleep mode can be used to determine energy savings due to power

DAC 2012 San Francisco, CA USA

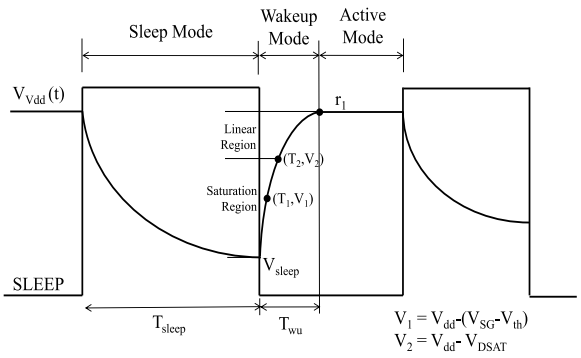

Figure 2: Typical timing instants and modes of operation in a power gating cycle

gating. In other words, some of the key design parameters have been captured in a single model.

The paper is organized as follows. Section 2 gives an overview of related work. Section 3 describes the equivalent circuit of a power-gated logic cluster. In Section 4, the semiempirical approach to estimation of wakeup time is described. The experimental results are presented in Section 5 to validate the model. Section 6 concludes the paper.

2. RELATED WORK

Several works have viewed design of power-gated circuits as an optimization problem of partitioning logic into clusters satisfying constraints of peak current, delay degradation, sleep transistor area, wakeup time and energy savings. Wakeup time estimation is fundamental to logic clustering algorithms proposed in [1], [2] and [6] that require constraints of peak current and wakeup time to be satisfied. In this context, a simple analytical model for estimation of wakeup time is useful especially when it needs to be determined iteratively during an optimization run for a number of candidate clusters. Run-time leakage reduction has been explored in [4] and [10], where only parts of the overall circuit are put to sleep during short periods of inactivity. Along with a high wakeup energy overhead a large wakeup delay in a cluster can result in reduced energy savings. Hence these two parameters have to be carefully considered in the design of power-gated circuits. The problem of wakeup time estimation arises in other scenarios as well. In [5] and [11], a need for wakeup latency estimation arises to quantify the effectiveness of proposed ground-bounce reducing techniques and intermediate strength power gating techniques respectively under a wakeup time constraint. In [14], Xu *et al.* have proposed numerical approaches for estimation of V_{Gnd} as a function of time in sleep mode. To extend the same method to wakeup mode, it is necessary to incorporate size dependent sleep transistor current characteristics. In this case an analytical model for V_{Gnd} would be highly desirable. This analysis is identically applicable to V_{Vdd} in a power-gated cluster with a header type of sleep transistor. Most works have used SPICE simulations or constant current source model [11] for sleep transistors to determine wakeup time. A simple expression for wakeup time based on a restrictive assumption of limited range of transistor operation was proposed in [13]. In this paper, the models are derived taking into account all the regions of sleep transistor operation necessary for applying them to large logic clusters consisting of high leakage cells.

3. POWER-GATED LOGIC CLUSTER MODEL

Models for subthreshold leakage current that capture its exponential behaviour with bias voltages at device level have been described in [12]. In [14], compact models for leakage current have been derived at gate and circuit levels in a hierarchical way. It was shown that the leakage current can be represented by a voltage controlled current source (VCCS) as in Fig. 1(b). In this work, we take a polynomial based approach to derive leakage current profile for the complete circuit. For each type of cell S_i and input pattern j, leakage current is determined at several voltages and the resulting profile is fitted with a polynomial of degree N in V_{Vdd} as given by

$$I_{leak}(S_i, j) = \sum_{k=0}^{N} b_k(S_i, j) V_{Vdd}^k \tag{1}$$

where $\{b_k(S_i, j)\}$ represents coefficients of the polynomial. We assume a standard-cell based design approach for implementation of the cluster. Therefore, the total static current for $n(S_i, j)$ occurrences of each cell and each input pattern is obtained as

$$I_{leak} = \sum_{i=0}^{P-1} \sum_{j=0}^{R_i-1} n(S_i, j) I_{leak}(S_i, j) \tag{2}$$

where P and R_i are number of types of cells and number of possible input combinations for cell S_i respectively. As an example, if a logic cluster is composed of a set $S = \{$NAND2, INV, NOR2, XOR2$\}$ of gates, then $P = 4$. For a 2-input NAND gate $R_i = 4$, whereas for an inverter, $R_i = 2$. For notational simplicity, the total leakage current profile of the logic cluster is represented by

$$I_{leak} = \sum_{i=0}^{N} b_i V_{Vdd}^i \tag{3}$$

in the rest of the paper. Equation (3) has the form of nonlinear resistance. The total capacitance of the logic cluster is derived as the sum of capacitances of all the inputs of all constituent standard cells.

$$C_L = \sum_{i=0}^{P-1} n(S_i) \sum_{l=0}^{R_i-1} C_{il} \tag{4}$$

4. VIRTUAL-VDD MODEL

4.1 Determination of Steady-State Virtual-Vdd Voltage

Consider the equivalent circuit model in Fig. 1(b). In the wakeup mode, the operating point on the I_{SD} vs. V_{SD} characteristics of sleep transistor moves from saturation region to linear region until V_{Vdd} reaches a steady-state value. The virtual supply node is said to be in steady state when $dV_{Vdd}/dt = 0$, i.e., when there are no changes in V_{Vdd} either on account of short-circuit currents due to changing logic states of internal nodes or due to charging effect. Let the current through the sleep transistor during wakeup and in non-saturation region be denoted by $I_{st,ns}$, the total leakage

current at the output of VCCS by I_{leak} and the capacitive load charging current by I_{load}. Then,

$$I_{st,ns} = I_{leak} + I_{load}. \quad (5)$$

The current through the sleep transistor in non-saturation region is given by the quadratic model

$$I_{st,ns}(t) = \frac{1}{R_{lin}} \left[(V_{dd} - V_{Vdd}(t)) - \frac{(V_{dd} - V_{Vdd}(t))^2}{2(V_{dd} - V_{th})} \right] \quad (6)$$

where R_{lin} is the resistance in linear region. The determination of R_{lin} is described in subsection 4.4. From (3), (6) and $I_{load} = C_L(dV_{Vdd}/dt)$, (5) becomes

$$\frac{dV_{Vdd}}{dt} = -\frac{1}{\tau} \sum_{i=0}^{N} c_i V_{Vdd}^i \quad (7)$$

where $\tau = R_{lin}C_L$ and $c_i = f_i(V_{dd}, R_{lin}, b_i, V_{th})$ are expressions derived from (3)-(6). To solve for V_{Vdd}, the Nth degree polynomial in (7) is reduced to a quadratic polynomial by least-squares approximation and is expressed in terms of its roots r_1 and r_2 as

$$\frac{dV_{Vdd}}{dt} = -\frac{1}{\tau}(V_{Vdd} - r_1)(V_{Vdd} - r_2). \quad (8)$$

Both r_1 and r_2 are steady state points of (8). One of the roots r_1 satisfying the interval of validity $V_{sleep} < r_1 < V_{dd}$, is determined to be the steady state Virtual-Vdd voltage. Here V_{sleep} denotes the value of V_{Vdd} at the wakeup transition. In a RC circuit the steady state as defined above is reached at $t = \infty$. However the error in assuming value of V_{Vdd} at onset of active mode to be r_1 is negligible as demonstrated in Section 5.

4.2 Wakeup Mode Virtual-Vdd Model

In order to obtain a model for $V_{Vdd}(t)$ in wakeup mode, the ordinary differential equation in (8) is solved in the non-saturation region and hence, is extended to saturation region by means of approximations. Let at time $t = 0$ the operating point move to non-saturation region so that $V_{Vdd}(0) = V_{initial}$. The solution of (8) satisfying the interval of validity and moving towards r_1 can be written as

$$[V_{Vdd}(t)]_{ns} = \frac{r_1 - r_2 K e^{-at}}{1 - K e^{-at}} \quad (9)$$

where $K = (V_{initial} - r_1)/(V_{initial} - r_2)$, $a = 1/A\tau$ and $A = 1/(r_1 - r_2)$. From Fig. 2, $V_{initial} = V_{dd} - V_{DSAT}$ where V_{DSAT} is the saturation voltage.

To extend the model to saturation region, the time instant $t = 0$ is moved to sleep-to-wakeup mode transition so that $V_{Vdd}(0) = V_{sleep}$. Let T_{wu} denote the wakeup time defined as the time taken for V_{Vdd} to evolve from V_{sleep} to $0.99r_1$. Further, let V_1 and V_2 be two voltage levels attained by V_{Vdd} at T_1 and T_2 respectively as shown in Fig. 2. The solution (9) does not represent V_{Vdd} in the saturation region, $V_{Vdd} < V_2$ accurately. Therefore corrections are applied to (5) in the first two segments as

$$I_{st}(t) = I_{leak} + I_{load} - \Delta I_0(t) + \Delta I_1(t). \quad (10)$$

In (10) the time instant $t = 0$ corresponds to sleep-to-wakeup mode transition and $V_{initial} = V_{sleep}$. Let U_T denote the time-shifted unit step function $u(t - T)$. We define

$$\Delta I_0(t) = I_0 \left[U_0 e^{-at} - U_{T_1} e^{-a(t-T_1)} \right] \quad (11)$$

$$\Delta I_1(t) = I_1 \left[U_{T_1} e^{-a(t-T_1)} - U_{T_2} e^{-a(t-T_2)} \right] \quad (12)$$

based on heuristics for I_0 and I_1 described in subsection 4.5. In the third interval (9) alone is satisfied and hence no correction is required. Using (9)-(12), the model for Virtual-Vdd in wakeup mode can be derived as

$$\begin{aligned} V_{Vdd}(t) = {} & \frac{r_1 - r_2 K e^{-at}}{1 - K e^{-at}} + A I_0 R_{lin} \left[U_0 \left(1 - e^{-at} \right) \right] \\ & - A R_{lin}(I_0 + I_1) \left[U_{T_1} \left(1 - e^{-a(t-T_1)} \right) \right] \\ & + A R_{lin} I_1 \left[U_{T_2} \left(1 - e^{-a(t-T_2)} \right) \right]. \end{aligned} \quad (13)$$

In compact form, (13) can be written as $Xe^{-2at} + Ye^{-at} + Z = 0$. The solution for t is given by

$$t = \frac{1}{a} \ln \left(\frac{2X}{-Y - \sqrt{Y^2 - 4XZ}} \right). \quad (14)$$

At $t = T_1$, $V_{Vdd} = V_{dd} - V_{SG} + V_{th}$ corresponding to the criterion $V_{SD} = V_{SG} - V_{th}$. Applying this condition and $V_{SG} = V_{dd}$, X, Y and Z are determined to be

$$\begin{cases} X = A R_{lin} K I_0, \\ Y = K(V_{th} - r_2) - (1 + K)X/K, \\ Z = A I_0 R_{lin} + r_1 - V_{th}. \end{cases} \quad (15)$$

Similarly for $t = T_2$, $V_{Vdd} = V_{dd} - V_{DSAT}$ corresponding to the condition $V_{SD} = V_{DSAT}$, which gives

$$\begin{cases} X = -A R_{lin} K [I_0(e^{aT_1} - 1) + I_1 e^{aT_1}], \\ Y = K(V_{dd} - V_{DSAT} - r_2) + (X/K) + A R_{lin} I_1 K, \\ Z = -A R_{lin} I_1 + r_1 - V_{dd} + V_{DSAT}. \end{cases} \quad (16)$$

For wakeup time T_{wu}, $V_{Vdd} = 0.99r_1$, which gives

$$\begin{cases} X = A R_{lin} K [-I_0(e^{aT_1} - 1) + I_1(e^{aT_2} - e^{aT_1})], \\ Y = K(0.99r_1 - r_2) - (X/K), \\ Z = 0.01r_1. \end{cases} \quad (17)$$

The values of r_1 and r_2 are unaffected due to ΔI_0 and ΔI_1 as (13) satisfies $V_{SD} = V_{DSAT}$ at $t = T_2$ as shown in Fig. 2. For clusters with $0.99r_1 \leq (V_{dd} - V_{DSAT})$, wakeup time $T_{wu} = T_2$ determined from (16) and with the condition $V_{Vdd} = 0.99r_1$.

4.3 Sleep Mode Virtual-Vdd Model

To calculate T_1, T_2 and T_{wu} using (15)-(17) it is necessary to determine V_{sleep}. If the cluster is in sleep state for a time interval T_{sleep}, then $V_{sleep} = V_{Vdd}(T_{sleep})$. It should be noted that for simplicity both mode transitions are assumed to occur at $t = 0$, so that the initial condition for sleep mode can be denoted by $V_{Vdd}(0)$ as for wakeup mode. In sleep mode, the sleep transistor is cut-off so that only a leakage current $I_{st,leak}$ flows through it. I_{load} in (5) is now a discharging current. Hence (5) for sleep mode can be written as

$$-C_L \frac{dV_{Vdd}}{dt} = -(b_0 - I_{st,leak}) - \sum_{i=1}^{N} b_i V_{Vdd}^i. \quad (18)$$

For each value of V_{Vdd}, the VCCS outputs a current given by (3). Therefore for each value of V_{Vdd} we infer that resistance of the circuit is given by $R_s(V_{Vdd}) = (V_{Vdd}/\sum_{i=0}^{N} b_i V_{Vdd}^i)$.

We refer to R_s as pseudo-resistance in the rest of the paper. Neglecting $I_{st,leak}$ and rewriting (18) similar to (7),

$$\frac{dV_{Vdd}}{dt} = -\frac{1}{R_s(V_{Vdd})C_L}\left[-R_s(V_{Vdd})\sum_{i=0}^{N} b_i V_{Vdd}^i\right]. \quad (19)$$

A numerical solution to (19) is of the form [14]

$$V_{Vdd,j+1} = V_{Vdd,j}e^{-\frac{\Delta t}{R_s(V_{Vdd,j})C_L}} \quad (20)$$

where j denotes a time interval in [0, T_{sleep}] of size Δt. To develop an approximation, we consider a heuristic for choice of R_s as explained in subsection 4.6. Denoting R_{sp} as the pseudo-resistance chosen by applying the heuristic, the model for Virtual-Vdd in sleep mode can be derived as

$$\frac{dV_{Vdd}}{dt} = -\frac{1}{R_{sp}C_L}\prod_{i=1}^{N}(V_{Vdd} - r_i^s) = 0 \quad (21)$$

where r_i^s represents roots of the polynomial in sleep context. Let r_1^s satisfy $r_1 < r_1^s < 0$. Then the approximate solution that moves towards r_1^s from its initial value is given by

$$V_{Vdd}(t) = r_1^s + e^{-\frac{t}{R_{sp}C_L}+K^s}. \quad (22)$$

At the end of active mode, the value of Virtual-Vdd satisfies $r_1 - \Delta V_{Vdd,max} \leq V_{Vdd} \leq r_1$ where $\Delta V_{Vdd,max}$ is the maximum degradation of V_{Vdd} due to dynamically changing inputs of logic cluster. In this work, it is assumed that the power-gated logic cluster remains in active mode for a duration long enough with appropriate input conditions that $V_{Vdd} = r_1$ at the end of active mode. This assumption is mostly true in circuits with adequate positive timing slack. Hence, applying the initial condition that $V_{Vdd}(0) = r_1$, we have

$$V_{Vdd}(t) = r_1^s + (r_1 - r_1^s)e^{-\frac{t}{R_{sp}C_L}}. \quad (23)$$

The value of V_{Vdd} at the end of sleep mode, V_{sleep}, is obtained by substituting $t = T_{sleep}$ in (23).

The energy savings E_s of the power-gated logic cluster in sleep mode with respect to an ungated cluster can be determined by

$$E_s = V_{dd}I_{leak}(V_{dd})T_{sleep} - \int_0^{T_{sleep}} V_{Vdd}I_{leak}(V_{Vdd})dt. \quad (24)$$

4.4 Determination of R_{lin}

To determine the resistance of sleep transistor in linear region, the method proposed in [7] for extraction of series-resistance (R_{sd}) of MOS device is followed. It is described here for completeness. Two operating points $(I_{SD}^{(1)}, V_{SG}^{(1)}, V_{th}^{(1)})$ and $(I_{SD}^{(2)}, V_{SG}^{(2)}, V_{th}^{(2)})$ with $V_{SD} = 0.05V$ are determined from I_{SD} vs. V_{SG} characteristics for a specific width W_{sp} of the transistor. All V_{SG} are chosen such that they satisfy constant mobility condition [7][12] while V_{th} is determined by g_m/I_D method. The drain current $I_{SD}^{(i)}$, for $i = 1, 2$, including the effects of R_{sd} is given by

$$I_{SD}^{(i)} = \mu C_{ox}\frac{W_{eff}}{L_{eff}}\left(V_{SG}^{(i)} - V_{th}^{(i)} - 0.5V_{SD}\right)\left(V_{SD} - R_{sd}I_{SD}^{(i)}\right). \quad (25)$$

Here μ is the constant carrier mobility, $C_{ox} = \epsilon_{ox}/t_{ox}$ is the oxide capacitance, W_{eff} and L_{eff} are effective width and channel length of sleep transistor. From the pair of equations (25), R_{sd} is determined. Further μ is determined from one of the equations of drain current in (25). Let R_{ch} denote the intrinsic channel resistance. Then $R_{lin} = R_{ch} + R_{sd}$. From [12],

$$R_{lin} = R_{sd} + \left(\frac{L_{eff}}{\mu C_{ox}W_{eff}(V_{SG} - V_{th} - 0.5V_{SD})}\right) \quad (26)$$

Table 1 shows linear region resistances for PMOS sleep transistors of different sizes in an industrial 65nm bulk CMOS technology library with nominal $V_{dd} = 1V$.

Table 1: Linear Region Resistance of PMOS Transistors (at 100°C, $L = 0.06\mu m$, $V_{SG} = 1V$, $V_{SD} = 0.05V$)

W (μm)	0.54	1.2	2.4	4.8	9.6	12
R_{lin} ($k\Omega$)	2.57	1.203	0.612	0.322	0.167	0.134

4.5 Heuristics for I_0 and I_1

Correction terms in (11) and (12) were applied in (10), to account for saturation region of sleep transistor operation. From (6) the current in saturation region is underestimated by $I_0 = I_{on,sat} - I_{st}(V_{Vdd} = V_{sleep})$ where $I_{on,sat}$ is the saturation drain current. Fig. 3 shows the variation of error in width-normalized estimated drain current $\frac{I_{error}}{W} = \frac{1}{W}(I_{SD} - I_{st})$ with V_{SD} where I_{st} is as determined from (6) for all V_{Vdd}. Similarly for I_1, we choose error in current corresponding to one of the values of V_{Vdd} in the interval $[(V_{dd} - V_{SG} + V_{th}), (V_{dd} - V_{DSAT})]$. From our experiments, we empirically choose $V_{SD} = 0.6V$, at which the error determined from Fig. 3 is $-I_1 = 0.174I_0$.

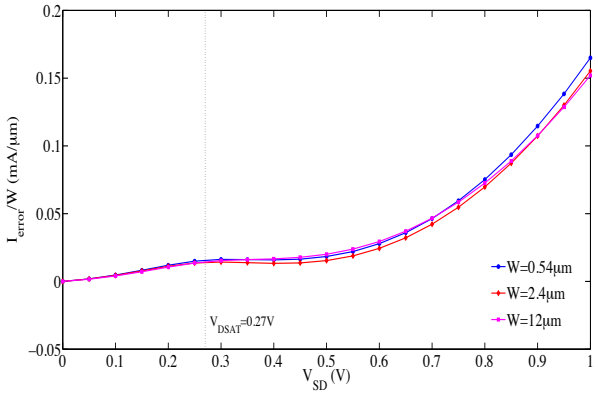

Figure 3: I_{error}/W vs. V_{SD} for 65nm PMOS transistors

4.6 Heuristic for R_{sp}

The voltage dependent pseudo-resistance changes as V_{Vdd} evolves with time according to (22). Hence it can be inferred that the time constant R_sC_L also varies with time. In our experiments, we have observed that in large logic clusters, the values of pseudo-resistance and its dynamic range are less than that for small logic clusters as leakage currents are higher in the former case. A typical variation of pseudo-resistance with V_{Vdd} is shown in Fig. 4 in the next section. The effect of a larger value of pseudo-resistance on V_{Vdd} is that it takes a longer time to change V_{Vdd} levels than with smaller values. Typically, higher values of pseudo-resistance determine V_{Vdd} after about 4 time constants of sleep time. Considering these observations, we choose R_{sp} as the pseudo-resistance at $V_{Vdd} = r_1$.

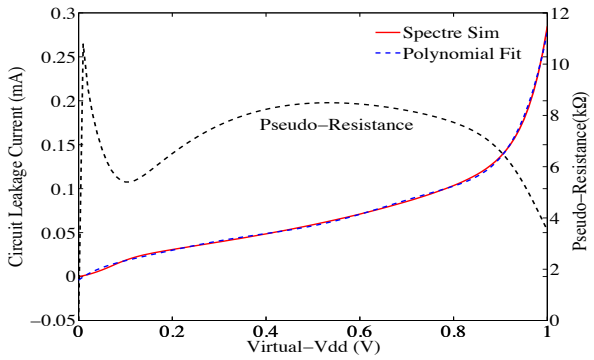

Figure 4: Leakage current and pseudo-resistance profile in c6288

5. EXPERIMENTAL RESULTS

The model was applied to ISCAS85 benchmark circuits [8] listed in Table 4 to validate the approximations proposed. The results were compared with simulations using Spectre circuit simulator of Cadence Virtuoso ICFB. Detailed results are reported for c7552, c6288, c2670 and c432 and a summary of results is provided for all circuits in Table 4.

The circuits were synthesized with two sets of logic gates, {nand2, nor2, xor2, and2, fa, ha, inv} in high-V_{th}(HVT) and {nand2, nor2, xor2, inv} in standard-V_{th}(SVT) process options of an industrial 65nm CMOS technology library. The two sets of circuits present wide variation in leakage current and total circuit capacitance for evaluation. For each logic gate, leakage currents were determined for supply voltage varying between 0 and 1V for all input patterns at an operating temperature of 100°C using Spectre. Each of these profiles were then fitted with polynomials of degree 7 using MATLAB. The maximum error between evaluated leakage current and simulated leakage current was less than 3% except near $V_{Vdd} = 0$, where absolute values of leakage current are negligible. Further, the leakage current profile of the complete circuit was determined by weighting the polynomials with number of occurrences in the gate netlist and adding them together to form I_{leak} in (3). A leakage current profile for c6288 is shown in Fig. 4. From this curve, pseudo-resistance is determined at each point in the Virtual-Vdd segment.

One set of Spectre simulations of high-V_{th} PMOS transistor is required for each technology library to determine threshold voltages, constant mobility, saturation voltage and saturation currents. To establish these parameters, I_{SD} vs. V_{SD} characteristics at $V_{SG} = 1V$ and I_{SD} vs. V_{SG} characteristics at $V_{SD} = 0.05V$ and $V_{SD} = 1V$ with $W_{sp} = 0.54\mu m$ were obtained using Spectre.

To compare wakeup time estimation using models with circuit simulations in Spectre, V_{dd} was set to 1V. Without loss of generality, all primary inputs of the circuit were set to logic 0. The evolution of Virtual-Vdd during wakeup and sleep modes in c7552 is shown in Fig. 5 and Fig. 6 respectively. In Table 2 and Table 3, the maximum voltage levels attained by Virtual-Vdd and the wakeup times with sleep transistors of different sizes are given. Table 4 shows average errors(μ_{error}) in estimation of the two quantities for all circuits considered in this work. The wakeup time is estimated by (15)-(17) within an average error margin of 16.3% for 22× variation in sleep transistor sizes. The steady-state Virtual-Vdd is determined within 1.8% on an average from the corresponding results of Spectre simulations. Further, a significant reduction in computation time is achieved for wakeup time estimation using the model compared to Spectre. For example, model calculations in c6288 using MATLAB took 21ms compared to 4 minutes in Spectre.

In logic clusters that do not satisfy wakeup dependency [2][6], short-circuit currents are generated due to changing logic states of internal nodes as V_{Vdd} increases towards r_1 in wakeup mode. They create the effect of altering effective resistance of the circuit and hence wakeup time. In other words, the accuracy of wakeup time estimation is reduced when the effects of short-circuit currents are not taken into account as is shown for c499 in Table 2 and Table 3. To address this problem it is necessary to model individual cells for short-circuit currents when both supply voltage and its rise time are varying. This is proposed for future work. The cluster definitions and sleep transistor widths considered in this work are not designed to satisfy wakeup dependency or meet a particular peak current constraint [6] as the problem of logic clustering is not addressed in this work.

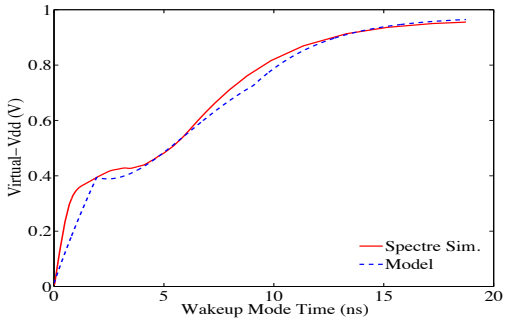

Figure 5: Virtual-Vdd in wakeup mode (W=1.2μm) in c7552

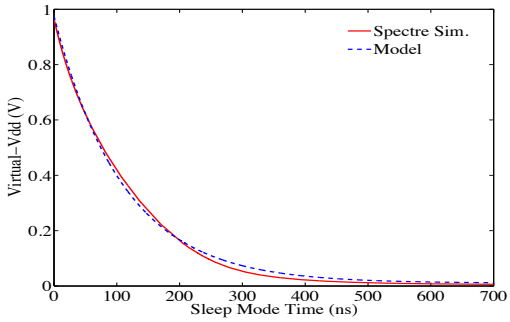

Figure 6: Virtual-Vdd in sleep mode (W=1.2μm) in c7552

6. CONCLUSION

In this paper, a semiempirical approach for estimation of wakeup time of a power-gated logic cluster that relies on only a few basic circuit parameters and one time SPICE level simulations per technology library was presented. A method to determine steady state Virtual-Vdd after wakeup as a function of sleep transistor size and leakage current profile was also described. This was fundamental to development of

Table 2: Maximum Virtual-Vdd after Wakeup and Wakeup Time (HVT Cells)

W (μm)	r_1 (V) Model [Eq. (8)] (Spectre)				Wakeup Time (ns) Model [Eq. (14),(17)] (Spectre)			
	c7552	c6288	c2670	c432	c7552	c6288	c2670	c432
0.54	0.95 (0.93)	0.96 (0.95)	0.97 (0.96)	0.99 (0.99)	40.69 (38.22)	43.24 (46.74)	15.55 (15.04)	4.57 (3.89)
1.2	0.98 (0.96)	0.98 (0.97)	0.99 (0.98)	0.99 (0.99)	18.94 (18.72)	20.27 (22.24)	7.38 (7.25)	2.20 (1.89)
2.4	0.99 (0.98)	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	9.64 (9.82)	10.35 (11.97)	3.79 (3.75)	1.14 (0.92)
4.8	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	5.08 (5.27)	5.46 (6.44)	2.00 (2.02)	0.60 (0.59)
9.6	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	2.64 (2.83)	2.84 (3.52)	1.05 (1.11)	0.32 (0.38)
12	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	0.99 (0.99)	2.13 (2.32)	2.29 (2.95)	0.85 (0.92)	0.26 (0.33)

Table 3: Maximum Virtual-Vdd after Wakeup and Wakeup Time (SVT Cells)

W (μm)	r_1 (V) Model [Eq. (8)] (Spectre)				Wakeup Time (ns) Model [Eq. (14),(17)] (Spectre)			
	c7552	c6288	c2670	c432	c7552	c6288	c2670	c432
0.54	0.73 (0.68)	0.68 (0.60)	0.77 (0.84)	0.97 (0.96)	36.66 (37.10)	62.79 (63.24)	18.06 (13.89)	4.65 (3.49)
1.2	0.85 (0.82)	0.82 (0.78)	0.88 (0.92)	0.99 (0.98)	20.90 (17.16)	34.59 (29.97)	8.17 (6.89)	2.21 (1.74)
2.4	0.91 (0.89)	0.90 (0.87)	0.93 (0.95)	0.99 (0.99)	10.01 (9.07)	16.65 (15.88)	4.01 (3.70)	1.14 (0.96)
4.8	0.95 (0.94)	0.94 (0.92)	0.96 (0.98)	0.99 (0.99)	5.19 (5.02)	8.45 (9.10)	2.07 (2.04)	0.60 (0.57)
9.6	0.97 (0.97)	0.97 (0.96)	0.98 (0.98)	0.99 (0.99)	2.65 (2.77)	4.3 (5.27)	1.06 (1.13)	0.32 (0.37)
12	0.98 (0.97)	0.98 (0.96)	0.98 (0.98)	0.99 (0.99)	2.13 (2.28)	3.45 (4.43)	0.86 (0.94)	0.26 (0.33)

Table 4: Average Relative Errors in Estimation of Maximum V_{Vdd} and Wakeup Time in ISCAS85 Benchmark Circuits

	C_L (pF)		Max. V_{Vdd} μ_{error} (%)		T_{wu} μ_{error} (%)	
Circuit	HVT	SVT	HVT	SVT	HVT	SVT
c7552	2.892	2.993	0.7	2.8	4.7	7.6
c6288	3.171	4.833	0.6	4.3	14.8	11.1
c5315	1.966	2.826	0.6	2.2	12.5	22.3
c3540	1.466	2.037	0.5	1.5	11.2	17.5
c2670	1.148	1.202	0.3	3.0	3.0	7.8
c1908	0.606	0.633	0.2	0.3	21.8	15.4
c499	0.601	0.478	0.2	0.1	24.6	33.7
c432	0.351	0.360	0.1	0.3	16.4	15.1
Mean			0.4	1.8	13.6	16.3

rest of the model. The model in sleep mode can be used to determine leakage energy savings in inactive states of the circuit. In other words, some of the key parameters used as optimization criteria for logic clustering have been captured in closed-form expressions. Our simulations and application of the model to ISCAS85 benchmark circuits with an industrial 65nm CMOS technology library show that on an average wakeup time can be estimated within an error margin of 16.3% over 22× variation in transistor sizes and 13× variation in circuit sizes with significant reduction in computational times compared to SPICE level circuit simulations.

7. REFERENCES

[1] A. Abdollahi, F. Fallah, and M. Pedram. An effective power mode transition technique in MTCMOS circuits. In *Proc. ACM/IEEE Des. Autom. Conf.*, pages 27–32, Anaheim, June 2005.

[2] M. Anis, S. Areibi, and M. Elmasry. Dynamic and leakage power reduction in MTCMOS circuits using an automated gate clustering technique. In *Proc. ACM/IEEE Des. Autom. Conf.*, pages 480–485, New Orleans, June 2002.

[3] S. Henzler. *Power Management of Digital Circuits in Deep Sub-Micron CMOS Technologies*, chapter 5. Springer, 2007.

[4] Z. Hu, A. Buyuktosunoglu, V. Srinivasan, V. Zyuban, H. Jacobson, and P. Bose. Microarchitectural techniques for power gating of execution units. In *Proc. Intl. Sym. Low Power Electronic Des.*, pages 32–37, Newport Beach, USA, August 2004.

[5] S. Kim, C. J. Choi, D. K. Jeong, S. V. Kosonocky, and S. B. Park. Reducing ground-bounce noise and stabilizing data-retention voltage of power gating structures. *IEEE Trans. Electron Devices*, 55(1):197–205, January 2008.

[6] Y. Lee, D.-K. Jeong, and T. Kim. Comprehensive analysis and control of design parameters for power gated circuits. *IEEE Trans. VLSI Syst.*, 19(3):494–498, March 2011.

[7] D.-W. Lin, M.-L. Cheng, S.-W. Wang, C.-C. Wu, and M.-J. Chen. A constant-mobility method to enable MOSFET series-resistance extraction. *IEEE Electron Device Lett.*, 28(12):1132–1134, December 2007.

[8] X. Lu. Layout and parasitic information for ISCAS circuits. `http://dropzone.tamu.edu/~xiang/iscas.html`.

[9] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits. *Proc. IEEE*, 91(2):305–327, June 2003.

[10] S. Roy, N. Ranganathan, and S. Katkoori. A framework for power-gating functional units in embedded microprocessors. *IEEE Trans. VLSI Syst.*, 17(11):1640–1649, November 2009.

[11] H. Singh, K. Agarwal, D. Sylvester, and K. J. Nowka. Enhanced leakage reduction techniques using intermediate strength power gating. *IEEE Trans. VLSI Syst.*, 15(11):1215–1224, November 2007.

[12] Y. Taur and T. Ning. *Fundamentals of Modern VLSI Devices*, chapter 4. Cambridge University Press, 2009.

[13] T.-D. Vivek, O. Sentieys, and S. Derrien. Wakeup time and wakeup energy estimation in power-gated logic clusters. In *Proc. 24th Intl. Conf. VLSI Des.*, pages 340–345, Chennai, January 2011.

[14] H. Xu, R. Vemuri, and W.-B. Jone. Dynamic characteristics of power gating during mode transition. *IEEE Trans. VLSI Syst.*, 19(2):237–249, February 2011.

SUPPLEMENTARY PAGES

S1. DETERMINATION OF TOTAL CIRCUIT CAPACITANCE

S1.1. Decoupling Capacitance

In (4), the total circuit capacitance C_L is defined to be the sum of input capacitances of all inputs of all constituient gates of the cluster. In physical implementations with CMOS process technologies, a decoupling capacitance (decap) C_D is generally included between the supply voltage rail (or Virtual-Vdd ring of the power-gated domain) and ground to suppress bounces on supply rails during switching of gate outputs. In the model described in this paper, a decap is not explicitly included. The total circuit capacitance including a decoupling capacitance can be determined to be $C_L + C_D$ since capacitances appear in parallel between Virtual-Vdd and ground.

S1.2. Dependence of Gate Capacitance on Inputs

The gate capacitances of MOS transistors are input dependent. At wakeup, as V_{Vdd} increases some of the logic gates switch to 'Logic 1' while the rest remain at 'Logic 0'. Gate inputs in the fan-out of gate outputs that switch to 'Logic 1' will present a higher output capacitance to the switching gate than to the driving gate remaining at 'Logic 0'. In logic clusters used in the paper, the outputs of each gate is determined from the primary inputs based on the gate function (NAND, XOR etc.) and hence appropriate value of capacitance obtained from SPICE level characterization of standard cell for each of its input is used to determine total capacitance in (4). Further gate terminal capacitances of all MOSFETs in standard cells include parasitic capacitances (fringe and overlap) referred to the gate terminal.

S1.3. Parasitic Capacitance Along Interconnect Lines

Parasitic capacitances along interconnect lines have been neglected in determining total circuit capacitance considering that in cluster based power-gated circuit design, independent clusters have a local distribution of interconnects unlike a distributed sleep transistor network (DSTN) based power gating.

S2. STEADY-STATE VIRTUAL-VDD

The conditions of validity for one of the roots r_1 can be intuitively explained to be $V_{sleep} < r_1 < V_{dd}$ as follows. Let at steady state the pseudo-resistance of VCCS be given by some $R_{ss} = V_{Vdd,ss}/I_{leak}(V_{Vdd,ss})$ where $V_{Vdd,ss}$ denotes Virtual-Vdd in steady state. Then the cicuit at that instant can be represented by Thévenin's equivalent resistance $R_{TH} = \frac{R_{ss}R_{lin}}{(R_{ss}+R_{lin})}$ and Thévenin's equivalent voltage $V_{TH} = \frac{R_{ss}V_{dd}}{(R_{ss}+R_{lin})}$ with total circuit capacitance C_L in series with R_{TH} and V_{TH}. Clearly, $V_{TH} < V_{dd}$. C_L is charged to V_{TH} in steady state which we determine to be r_1 as a solution of (8). For non-zero C_L the inference that $V_{sleep} < r_1$ is trivial.

S3. WAKEUP MODE VIRTUAL-VDD MODEL

In (6), V_{th} corresponds to threshold voltage when the transistor is operating in linear region as determined from subsection 4.4.

To derive (8), dV_{Vdd}/dt obtained from evaluation of RHS of (7) for a sweep of V_{Vdd} is fitted with a quadratic polynomial in least squares sense using the MATLAB function 'polyfit'. Further r_1 and r_2 are obtained as roots of the quadratic polynomial on the RHS of (8). Equation (9) can be derived from (8) by separation of variables and partial fraction expansion as

$$dV_{Vdd}\left(\frac{A}{V_{Vdd}-r_1}+\frac{B}{V_{Vdd}-r_2}\right) = -\frac{1}{\tau}dt.$$

Equations (10) and (13) denote I_{st} and V_{Vdd} represented by piecewise continuous functions in three intervals: two in saturation region and one in non-saturation region of sleep transistor operation.

The wakeup time T_{wu} is given by (17) under the assumption that $0.99r_1 > (V_{dd} - V_{DSAT})$. For clusters with $0.99r_1 \leq (V_{dd} - V_{DSAT})$, wakeup time $T_{wu} = T_2$ determined from (16) and with the condition $V_{Vdd} = 0.99r_1$.

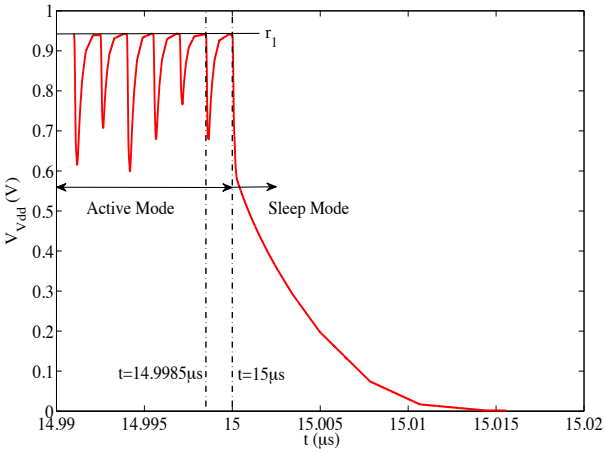

Figure 7: Virtual-Vdd in Active and Sleep Modes

S4. SLEEP MODE VIRTUAL-VDD MODEL

In this section we justify the assumption that at the end of active mode $V_{Vdd} = r_1$ with an experiment. Let $\Delta V_{Vdd,max}$ denote maximum degradation of V_{Vdd} due to dynamically changing inputs of logic cluster. Further, let t_i denote the time instant of end of cycle i in active mode. Clearly, $V_{Vdd}(t_i)$ satisfies $V_{Vdd}(t_{i-1}) - \Delta V_{Vdd,max} \leq V_{Vdd}(t_i) \leq r_1$. We consider the conditions under which $V_{Vdd}(t_{i-1}) = r_1$.

A logic cluster was synthesized for a maximum path delay of 1.0ns and was power-gated by a header type of sleep transistor. In practice, the size of the sleep transistor is chosen for a fixed performance loss or $\Delta V_{Vdd,max}$ in active mode. The circuit was simulated with Spectre circuit simulator with random inputs applied to the circuit at a clock period of 1.5ns, i.e., with a positive slack of 0.5ns in active mode. The variation of Virtual-Vdd voltage in active and sleep modes is shown in Fig. 7. It can be seen that for a maximum duration of path delay, V_{Vdd} degrades by about $\Delta V_{Vdd,max} = 0.3V$ and at the end of active mode time slot of 1.5ns, V_{Vdd} attains a value of r_1. With a sufficient and constant clock cycle period $T = t_i - t_{i-1}$, $V_{Vdd}(t_{i-1}) = r_1$

for all i. The assumption of sufficient positive slack holds for low power, low performance circuits. Therefore $V_{Vdd} = r_1$ can be specified as the initial value of V_{Vdd} in sleep mode in (22).

S5. EXPERIMENTAL RESULTS

To apply the model and to perform simulations on ISCAS85 benchmark circuits all primary inputs were assigned 'Logic 0'. In practice this input combination may not result in minimum leakage current. However the model for estimation of wakeup time described in the paper applies identically to all patterns of primary inputs.

The ISCAS85 benchmark circuits listed in Table 4 were synthesized with both HVT and SVT cells of an industrial 65nm bulk CMOS technology library. Since threshold voltage of MOSFET devices in SVT cells is less than that of HVT cells, the total leakage current in SVT cell implementations of ISCAS85 circuits is higher than that of HVT cell implementations resulting in a lower R_{ss} as defined in Section S2. Hence a lower V_{TH} or r_1 is obtained. This observation is reflected in the results shown in Table 2 and Table 3.

In this work MATLAB was used to evaluate model parameters and hence wakeup time. Alternatively other tools may be used for computations. As an example, the same model, when implemented in a scripting language like Tcl, can be efficiently integrated with standard IC design and analysis flows that use static timing and power analysis tools.

S6. APPLICATIONS

S6.1. Wakeup Energy Estimation

As an extension to estimation of wakeup time presented in the paper, it is possible to determine wakeup energy (E_{wu}).

Wakeup energy is an energy overhead due to sleep-to-wakeup mode transition in a power gating cycle. It is required to determine breakeven energy and hence minimum sleep time for a power-gated logic cluster. Wakeup energy is given by $E_{wu} = \int_0^{T_{wu}} V_{dd} I_{st} dt$ where I_{st} is obtained from (10) and (13) and T_{wu} from (17). It should be noted that, short circuit currents (I_{sc}) that are generated in the internal nodes during wakeup mode are neglected in (5). In this work I_0 and I_1, which are assumed to be constants based on heuristics developed in subsection 4.5, must be replaced with time and V_{Vdd} dependent models. Modeling short circuit currents in logic gates when both supply voltage and its rise time are varying is proposed for future work.

S6.2. Scheduling Power-Gated Clusters

The model for wakeup time estimation presented in the paper may be applied in scheduling of power-gated logic clusters as part of a larger optimization problem. Consider a combinational circuit **C**. Let $C_i, i = 1, 2, ..., N$ denote N logic clusters obtained by partitioning **C** such that they satisfy constraints of minimum sleep transistor area, peak current, maximum delay degradation and minimum wakeup time. The optimization problem referred to wakeup time constraint is stated as follows. Let $T_{wu,i}$ denote the wakeup time of logic cluster C_i and $T_{wu,max}$ the maximum acceptable wakeup time of the overall circuit **C**. Then,

$$max\left(\sum_{j=1}^{P} T_{wu,j}, \sum_{k=P+1}^{Q} T_{wu,k}, ..., \sum_{l=R+1}^{N} T_{wu,l}\right) \leq T_{wu,max}$$

for some P,Q,R,... such that $P \geq 1$, $Q \geq P + 1$,.... Hence a wakeup schedule for the N logic clusters may be derived. The model presented in the paper may be used to determine each $T_{wu,i}$ during the optimization run.