# Long-Term Power Minimization of Dual-$V_T$ CMOS Circuits

*Suhwan Kim, Youngsoo Shin, Stephen Kosonocky, and Wei Hwang*

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

## ABSTRACT

In this paper, we define *Long-Term* power dissipation in which the effect of the system-level power management on the total power dissipation of a given circuit is considered. Then, we present a novel design methodology to minimize the *Long-Term* power dissipation of a circuit used along with dual-threshold voltage selection and voltage scaling. In simulation on 16-bit carry lookahead adders (CLAs), the proposed approach can reduce up to 80% and 25% of the total power dissipation along with clock- and power-gating, respectively.

## I. INTRODUCTION

With the recent trend toward portable system-on-chip (SOC) for communication and computing, power dissipation has become a critical design constraint. Most hand-held devices are characterized by intermittent operations with long period of idle, in which the leakage power should be considered as a major component of the total power dissipation in CMOS circuits.

Recently, to meet target clock frequency requirement of performance-critical blocks and to minimize overall leakage power dissipation of the blocks, the use of dual-threshold voltage transistors (dual-$V_T$) has been studied [1], [2], [3], [4]. However, the previous works have not considered the effect of system-level dynamic power management on the dual-$V_T$ selection and voltage scaling of a given circuit, despite its significance.

In this paper, we define *Long-Term* power dissipation in which the effect of the system-level dynamic power management on the total power dissipation of a given circuit is considered. Then, we present a novel design methodology to minimize the power dissipation of dual-$V_T$ CMOS circuits for battery-operated portable system. In order to minimize the *Long-Term* power dissipation within performance constraint, we find an optimal mapping of circuit components to different threshold voltages, and assign an optimal supply voltage to entire circuit. The algorithm simultaneously optimizes the percentage of $V_{T,low}$ and $V_{DD}$, considering the long-term effect on the total power dissipation. We then show how dual-threshold voltage selection and voltage scaling can be considered with clock- or power-gating for reducing *Long-Term* power dissipation.

The remainder of the paper is organized as follows. The motivation of our works is described in Section II and the *Long-Term* power dissipation, which is the optimization objective, is defined in Section III. Section IV describes our algorithm

that targets the reduction of the *Long-Term* power dissipation in circuits during active and sleep state. Section V presents simulation results from 16-bit carry lookahead adders (CLAs) whose clock and power are gated in sleep state, respectively. The CLAs are designed using a 0.13 $\mu$m dual-threshold CMOS technology. Our contributions are summarized in Section VI.

## II. DYNAMIC POWER MANAGEMENT

Most hand-held devices such as portable computers, cellular phones, and PDAs spend a small fraction of time performing useful computation. The rest of the time is spent idling between user requests. However, when the bursts of useful computations are demanded, the faster throughput, the better. This characterizes the burst throughput mode of computation as shown in Figure 1.
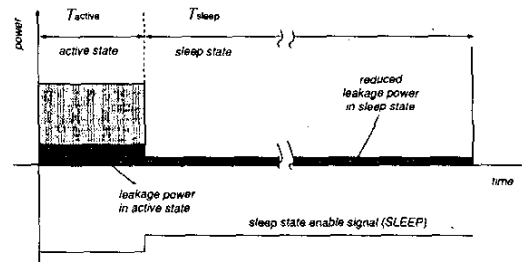


Fig. 1. Power management sequences in burst throughput mode.

Dynamic power management - which refers to the selective shut-off or slow-down of system components that are idle or underutilized - is known as one of the most successful low power design techniques in such systems [5]. One of the most popular methods used to implement the dynamic power management is *clock gating*: whenever a functional unit becomes idle, its clock signal is stopped, preventing power consumption caused by spurious switching. Many commercial microprocessor implement clock gating to reduce power dissipation. However, the processor still dissipates leakage power, which is not eliminated by clock gating.

In order to deal with the leakage problem, *power gating* such as multiple threshold voltage CMOS (MTCMOS) circuit has been considered [6]. The gates of logic core are composed of low threshold-voltage ($V_{T,low}$) MOSFETs to enhance speed at low voltages. A high threshold-voltage ($V_{T,high}$) MOSFET connects the power supply to the logic core. This MOSFET, what is called as *header*, is controlled by a sleep signal, which
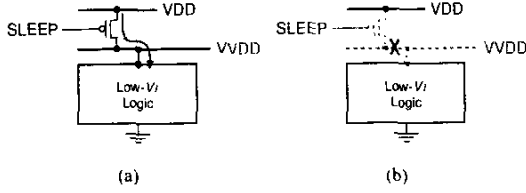
Fig. 2. MTCMOS in (a) active or (b) sleep state.

switches between the active and sleep state. In active state, as shown in Figure 2 (a), the MOSFET supplies power to the logic core. In the sleep state, as shown in Figure 2 (b), the MOS-FET suspends power supply to cut the standby leakage current. However, the large inserted MOSFETs increase the area and delay.

### III. Long-Term POWER DISSIPATION

In this section, we define Long-Term power dissipation in which the effect of dynamic power dissipation on the total power consumption of a given circuit is considered as shown in Equation (1).

$$P_{Long-Term} = \eta \cdot P_{active} + (1 - \eta) \cdot P_{sleep} \qquad (1)$$

where $P_{active}$ and $P_{sleep}$ are the power consumption in the active and sleep states and $\eta$ is the active-time ratio. The $\eta$ is defined as $T_{active}/(T_{active} + T_{sleep})$, where $T_{active}$ and $T_{sleep}$ shown in Figure 1 are the times of the system operated in active and sleep state, respectively.

$$
\begin{aligned}
P_{active} &= P_{sw} + P_{sc} + P_{leakage\_active} \qquad (2) \\
&= \alpha \cdot C_L \cdot V_{DD\_active}^2 \cdot f_{clk} \\
&+ I_{sc} \cdot V_{DD\_active} \\
&+ I_{leakage\_active} \cdot V_{DD\_active}
\end{aligned}
$$

In active state, the power dissipation in CMOS circuits is given by Equation (2). The dynamic switching power $P_{sw}$ has been traditionally considered to be the dominant component, where $\alpha$ is the expected number of transitions per cycle, $C_L$ is the capacitive load, and $f_{clk}$ is the clock frequency. The short circuit power $P_{sc}$ is due to the existence of a direct current path between $V_{DD}$ and ground during the short period of time when both the PMOS and NMOS transistors are simultaneously turned on.

$$
\begin{aligned}
P_{sleep} &= P_{leakage\_sleep} \qquad (3) \\
&= I_{leakage\_sleep} \cdot V_{DD\_sleep}
\end{aligned}
$$

In sleep state, as shown in Equation (3), the power dissipation is due to the leakage current caused by the stored charge in drain junctions leaking away, or gate oxide tunneling current, and due to sub-threshold current.

Meanwhile, clock and power-gating described in the previous section have different power and timing characteristics

that affects the Long-Term power dissipation of Equation (1). Power dissipation in sleep state can also be defined as follows.

$$P_{sleep} = \left(\frac{\gamma}{1-\gamma}\right) \cdot P_{active} \qquad (4)$$

where $P_{active}$ and $P_{sleep}$ are the power consumption in the active and sleep states and $\gamma$ is the power-down efficiency of sleep state defined as $P_{sleep}/(P_{active} + P_{sleep})$ in Figure 1.

### IV. OPTIMIZATION PROCEDURE

In this section, our main objective is to minimize Long-Term power dissipation of Equation (1) through dual-threshold voltage selection and supply voltage. More formally, our optimization problem can be stated as follows: for a given a combinational circuit with proper timing constraint, range of supply voltage $[V_{DD,min}, V_{DD,max}]$, and active-time ratio $\eta$, derive $V_{DD}$ and a mapping $g_i \rightarrow \{V_{T,low}, V_{T,high}\}$ ($g_i$ denotes a gate in critical path, $V_{T,low}$ and $V_{T,high}$ are low- and high-threshold voltages, respectively); such that timing constraint is satisfied and $P_{Long-Term}$ is minimized.

The inherent nature of the optimization problem indicates that it is very unlikely that we can find an optimal solution ($V_{DD}$ and mapping) with a polynomial time complexity. In order to solve the problem in efficient way, we propose a heuristic algorithm, whose overall procedure is shown in Figure 3 between L1 and L6. We assume that all gates are initially mapped to $V_{T,high}$ and the timing constraint is equal to the critical path delay for this initial mapping with the highest supply voltage. that is $V_{DD} = V_{DD,max}$.

At each iteration of the while loop L1, we generate a mapping L2, denoted by $M_i$, which assigns either $V_{T,low}$ or $V_{T,high}$ to each gate. This is followed by power optimization procedure L4, which finds the lowest supply voltage for a specific mapping within the performance constraint. Among the resulting mappings with their associated supply voltages, we select the one that minimizes the Long-Term power dissipation $P_{Long-Term}$.

Each mapping is generated starting from the previously generated one as shown in the procedure GENMAPPING. It finds a critical path L9, and if there are gates in the critical path with $V_{T,high}$, all of those gates are assigned $V_{T,low}$ L12 that leads to a new mapping. This gives the critical path a time margin or slack, meaning that we have a chance to reduce the supply voltage until timing constraint is just met, which contributes to reducing both power components ($P_{active}$ and $P_{sleep}$) in (1). If all gates in the critical path are already mapped to $V_{T,low}$ L11, supply voltage is reduced that leads to a potentially new critical path. In other words, the critical path at the specific supply voltage may not be critical if we reduce the supply voltage. If we cannot find a new mapping within the performance constraint and the range of supply voltage, we stop the entire process.

### V. SIMULATION RESULTS

The Long-Term power optimization algorithm described in Section IV is evaluated with several benchmark circuits de-

Fig. 4. Schematic of 16-CLA.



(a)                    (b)

Fig. 5. Critical paths of 16-bit CLA. (a) when all of transistors are mapped to $V_{T,high}$ with the supply voltage of 0.9V and (b) when all of transistors are mapped to $V_{T,low}$ with the supply voltage of 0.775V.



Fig. 6. Each step of optimization procedure as the percentage of transistors mapped to $V_{T,low}$ gradually increase and the supply voltage of the circuit decreases within performance constraint.

signed in 0.13$\mu$m Dual-$V_T$ bulk CMOS technology. The randomly generated input patterns and transistor sizes are not changed during optimization. The timing constraint for each circuit is assumed to be the critical path delay when the threshold voltage of all transistors are $V_{T,high}$ and supply voltage is the maximum. In order to show how our optimization procedure works we use 16-bit carry lookahead adders (CLAs) whose clock and power are gated in sleep state, respectively.

The schematic of 16-bit CLA is shown in Figure 4. The triangles are buffers. The rectangular box is the basic cell and it is composed of one AND gate and one AOI gate.

Figure 5 (a) shows the critical path of the original 16-bit CLA in which all of transistors are mapped to $V_{T,high}$, when the supply voltage for an entire circuit is 0.9V. Figure 5 (b) shows the critical path of the 16-bit CLA in which all of transistors are mapped to $V_{T,high}$, when the supply voltage for an entire circuit is 0.775V.

Figure 6 shows each step of the optimization procedure described in the previous section. During the optimization of the circuit, the percentage of transistors mapped to $V_{T,low}$ gradually increases and the supply voltage of the circuit decreases within performance constraint. At the beginning of optimiza-
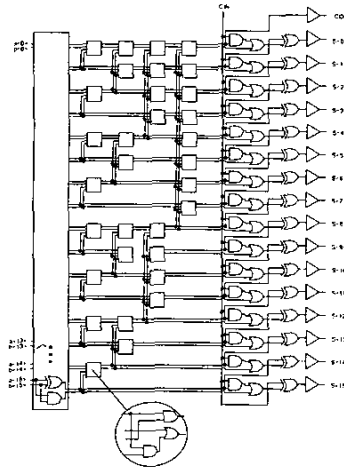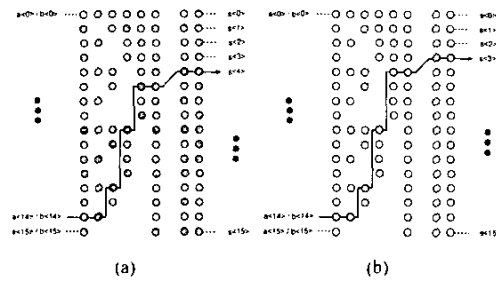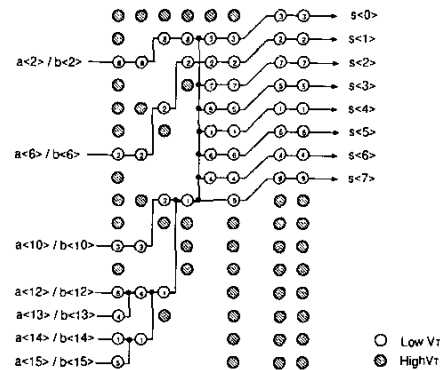
tion procedure, all of transistors in the 16-bit CLA are assigned to $V_{T,high}$ and the supply voltage is 0.9V. Then, all of transistors in its worst-case critical path are replaced to $V_{T,low}$ as marked

"1" in Figure 6. As a result, the supply voltage can be reduced down to 0.85V within performance constraint. Finally, *Long-Term* power dissipation of the circuit is obtained, considering the active-time ratio ($\eta$) of target system and power-down efficiency ($\gamma$) of sleep state. This optimization step is repeated until we cannot find a new mapping within the performance constraint and the range of supply voltage.
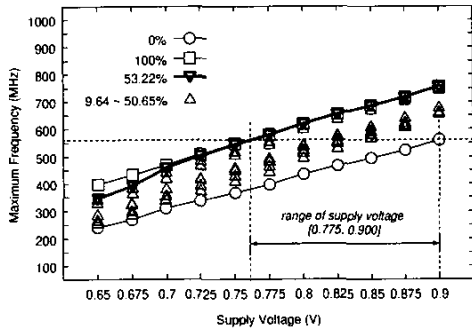


Fig. 7. Maximum frequencies of 16-CLAs optimized with different percentage of the transistors mapped to $V_{T,low}$ as a function of supply voltage.

Figure 7 shows maximum frequencies of 16-bit CLAs optimized with the different percentage of transistors mapped to $V_{T,low}$ as a function of supply voltage. Within performance constraint, the maximum frequencies of the 16-bit CLA in which only 53.22% of transistors are mapped to $V_{T,low}$ is equal to those of the 16-bit CLA in which all of transistors are mapped to $V_{T,low}$.
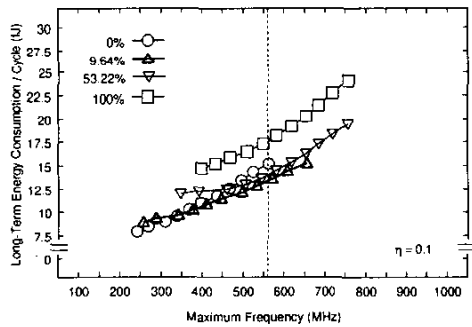


Fig. 8. *Long-Term* energy consumption per cycle of 16-bit CLA that uses clock gating in sleep state as a function of maximum frequency when $\eta = 0.1$.

Figure 8 and 9 show the *Long-Term* energy consumption per cycle of 16-bit CLAs that use clock gating in sleep state as a function of maximum frequency when the percentage of transistors mapped to $V_{T,low}$ varies as 0%, 9.64%, 53.22%, and 100%. When the $\eta$ is changed from 0.1 to 0.01, the optimal percentage of transistors mapped to $V_{T,low}$ varies from 52% to 9.64%.

Figure 10 shows the *Long-Term* energy reduction per cycle of 16-bit CLAs in which the percentage of transistors mapped
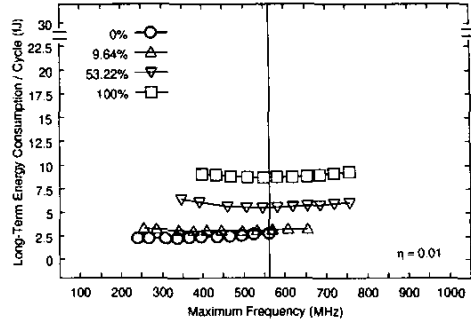


Fig. 9. *Long-Term* energy consumption per cycle of 16-bit CLA that uses clock gating in sleep state as a function of maximum frequency when $\eta = 0.01$.
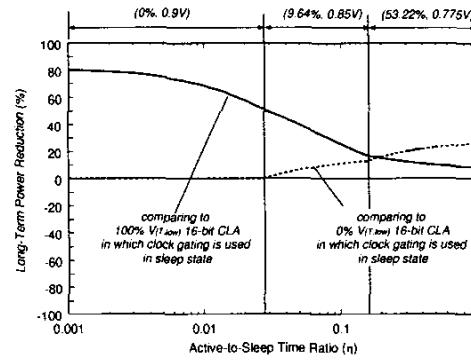


Fig. 10. *Long-Term* energy reduction per cycle of 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized within performance constraint and clock gating is used during sleep state.

to $V_{T,low}$ is optimized within performance constraint and clock gating is used during sleep state, comparing with 16-bit CLAs in which all of transistors are mapped to $V_{T,low}$ or $V_{T,high}$ and clock gating is also used during sleep state.
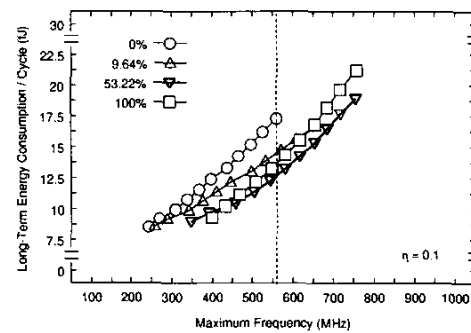


Fig. 11. *Long-Term* energy consumption per cycle of 16-bit CLA that uses power gating in sleep state as a function of maximum frequency when $\eta = 0.1$.

Figure 11 and 12 show the *Long-Term* energy consumption per cycle of 16-bit CLAs that use clock gating in sleep state as a function of maximum frequency when the percentage of
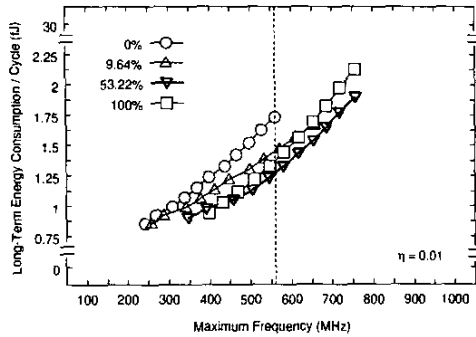
Fig. 12. *Long-Term* energy consumption per cycle of 16-bit CLA that uses power gating in sleep state as a function of maximum frequency when $\eta = 0.01$.

transistors mapped to $V_{T,low}$ varies as 0%, 9.64%, 53.22%, and 100%. Even though the $\eta$ is changed from 0.1 to 0.01, the optimal percentage of transistors mapped to $V_{T,low}$ is the same as 52%.
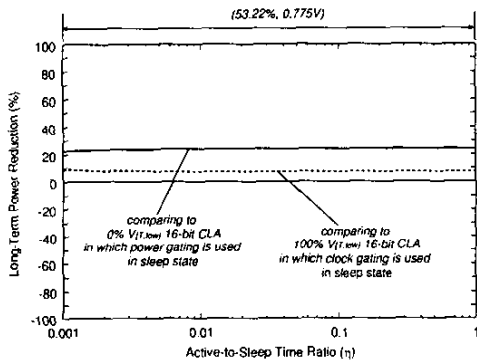


Fig. 13. *Long-Term* power reduction of 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized within performance constraint and power gating is used during sleep state.

Figure 13 the relative *Long-Term* power reduction of 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized within performance constraint and power gating is used during sleep state, comparing with 16-bit CLAs in which all of transistors are mapped to $V_{T,low}$ or $V_{T,high}$ and power gating is also used during sleep state.

Figure 14 the relative *Long-Term* power reduction of 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized within performance constraint and power gating is used during sleep state, compared to 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized within performance constraint and clock gating is used during sleep state.

When the $\eta$ is small, the circuits with optimized Dual-$V_T$ mapping and power gating is more attractive than the circuits with optimized Dual-$V_T$ mapping and clock gating, in terms of terms of *Long-Term* power minimization. To keep its original
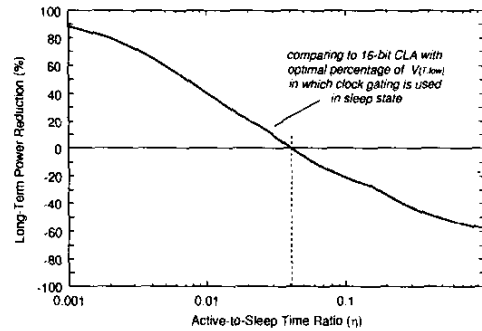


Fig. 14. *Long-Term* power reduction of 16-bit CLAs with optimized percentage of transistors and power gating for sleep state, compared to 16-bit CLAs with optimized percentage of transistors and clock gating for sleep state.

performance, however, the 16-bit CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized and power gating is used during sleep required higher supply voltages than that of 16-CLAs in which the percentage of transistors mapped to $V_{T,low}$ is optimized and clock gating is used during sleep As a result, the CLAs with clock gating is more attractive when the $\eta$ become large.

## VI. CONCLUSION

We have successfully defined *Long-Term* power dissipation in which the effect of the system-level dynamic power management on the total power dissipation of a given circuit is considered. We also presented a novel design methodology to minimize the *Long-Term* power dissipation of a circuit used along with dual-threshold voltage selection and voltage scaling. 16-bit carry lookahead adders (CLAs) designed and optimized in a 0.13$\mu$m dual-threshold CMOS have been simulated. The results show that the total power dissipation of the CLAs can be reduced up to 80% and 25% with clock- and power-gating, respectively.

## REFERENCES

[1] Z. Chen, C. Diaz, J. D. Plummer, M. Cao, and W. Greene, "0.18 um dual Vt MOSFET process and energy-delay measurement," in *Proceedings of IEEE International Electron Devices Meeting*, 1996, pp. 851–854.

[2] Q. Wang and S. Vrudhula, "Efficient procedures for minimizing the standby power in dual $V_T$ CMOS circuits," in *Proceedings of International Workshop on Power and Timing Modeling, Optimization and Simulation*, Oct. 1998.

[3] Q. Wang and S. Vrudhula, "Static power optimization of deep submicron cmos circuits for dual $V_T$ technology," in *Proceedings of the International Conference on Computer Aided Design*, 1998, pp. 490–494.

[4] P. Pant, R. K. Roy, and A. Chatterjee, "Dual-threshold voltage assignment with transistor sizing for low power CMOS circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 9, no. 2, pp. 390–394, Apr. 2001.

[5] V. Tiwari, R. Donnelly, S. Malik, and R. Gonzalez, "Dynamic power management for microprocessors: a case study," in *Proceedings Tenth International Conference on VLSI Design*, Jan. 1997, pp. 185–192.

[6] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamda, "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS," *IEEE Journal of Sold-State Circuits*, vol. SC-30, no. 8, pp. 847–854, Aug. 1995.