Low-Power CMOS Design through $V_{TH}$ Control and Low-Swing Circuits

Takayasu Sakurai*, Hiroshi Kawaguchi* and Tadahiro Kuroda**

*) Institute of Industrial Science, Univ. of Tokyo, 7-22-1, Roppongi, Minato-ku, Tokyo, 106 Japan
E-mail:tsakurai@iis.u-tokyo.ac.jp
***) Microelectronics Engineering Lab., Toshiba Corporation

Abstract
This paper describes some of the circuit level techniques for low-power CMOS designs. $V_{TH}$ control circuits are necessary for achieving low-threshold voltage in high-speed low-voltage applications. As for the low-swing circuit techniques, applications to a clock system, logic part, and I/O's are discussed.

1. Introduction
CMOS power dissipation and delay are given by

$$P = p_{L} \cdot C_{L} \cdot V_{S} \cdot V_{DD} \cdot f_{C} + l_{S} \cdot 10^{-3} \cdot V_{DD}.$$ (1)

The first term in (1) represents dynamic power dissipation due to charging and discharging of the load capacitance, where $p_{L}$ is the switching probability, $C_{L}$ is the load capacitance, $V_{S}$ is the voltage swing of a signal, and $f_{C}$ is the clock frequency. The second term is the threshold voltage term and $S$ is typically about 100mV/decade.

Figure 1 shows the plot for power and delay assuming 0.5um design rule. As seen from the figure, lowering $V_{DD}$ is effective in decreasing power but delay increases. Fig. 1(b) shows equi-delay curves and the delay can be maintained if the $V_{TH}$ is lowered as $V_{DD}$ is reduced. Lowering $V_{TH}$, however, increases subthreshold leakage. In order to cope with this problem, $V_{TH}$ control schemes have been proposed which are covered in Section 2.

In most cases, $V_{S}$ in (1) is the same as $V_{DD}$, but in low-swing circuits $V_{S}$ is smaller than $V_{DD}$. As seen from Eq.(1), reducing $V_{S}$ can be one promising way to decrease power consumption. As for the low-swing circuit techniques, applications to a clock system, logic part, and I/O's are discussed in Section 3, 4, and 5, respectively.

2. $V_{TH}$ control techniques
To maintain throughput while lowering supply voltage to decrease power consumption, it is effective to lower the threshold voltage of MOSFET's. There are, however, issues associated with low $V_{TH}$ in low $V_{DD}$ environments.

First, delay fluctuates intolerably with $V_{TH}$ fluctuation in low $V_{DD}$ regime. For example, delay increase by 3 times for $AV_{TH} = +0.15V$ at $V_{DD}$ of 1V. The second issue is the subthreshold leakage increase. The leakage increases by 10 times for every $AV_{TH}$ of -0.1V. The third problem is the inability for $I_{DQ}$ test. $I_{DQ}$ test is necessary to screen out LST's with defects and micro-shots which develop to a failure in a long run.

In order to cope with these issues, $V_{TH}$ control techniques have been proposed which are summarized in Table 1.

| TABLE 1. Multi-Threshold $V_{TH}$ CMOS[3,4], MTCMOS in short, tries to decrease the subthreshold leak in standby mode by inserting high $V_{TH}$ MOSFET in series to normal circuitry. The high-$V_{TH}$ device is turned off in standby mode and completely cut-off the leakage path. The drawback is the large inserted MOSFET which increases area and delay. While the MTCMOS can solve only the standby leakage problem, the Variable Threshold CMOS[5-9] (VT CMOS) can solve all the three problems. It dynamically varies $V_{TH}$ through substrate-bias, $V_{BB}$. Typically, $V_{BB}$ is controlled so as to compensate $V_{TH}$ fluctuations in the active mode, while in the standby mode and in the $I_{DQ}$ testing, deep $V_{BB}$ is applied to increase $V_{TH}$ and cut off the subthreshold leakage current. The idea to control the $V_{BB}$ so as to minimize the subthreshold leakage under the condition that a representative circuit shows sufficient speed was also proposed (Frequency adaptive Threshold CMOS, FTCMOS[10]).

The Elastic $V_{TH}$ CMOS[11], EVTCMOS in short, controls both $V_{DD}$ and $V_{BB}$ such that when $V_{DD}$ is lowered VBB becomes much deeper to raise $V_{TH}$ and further reduce power dissipation. Note that internal $V_{DD}$ and VSS are provided by source-follower n- and p- transistors, respectively, whose gate voltages are controlled. In order to control the internal power supply voltage independent from the power current, the source-follower transistors should operate near the threshold. This requires very large transistors.

In VTCMOS, it has been experimentally evaluated that the number of substrate (well) contacts can be greatly reduced in low voltage environments [7-9]. Using a phase-locked loop and an SRAM in a VTCMOS gate-array [8], the substrate noise influence has been shown to be negligible even with 1/400 of the contact frequency compared with the conventional gate-array. A DCT (Discrete Cosine Transform) macro made with the VTCMOS [7] has also been manufactured with substrate- and well- contacts only at the periphery of the macro and it worked without problems realizing more than one order of magnitude smaller power dissipation than a DCT macro in the conventional CMOS design.

3. Low-swing circuit for clock system
The four pie charts in Fig.2 shows power distribution in VLST's. As seen from the charts, the power distribution of VLST's differs from product to product. However, it is interesting to note that a clock system and a logic part itself consume almost the same power in various chips, and the clock system consumes 20% to 45% of the total chip power. One of the reasons for this large power consumption of the clock system is that the transition ratio of the clock net is

Permission to make digital/hard copy of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

\*\*1997 ACM 0-89791-903-3/97/08...53.50
one while that of the ordinary logic is about one third on average.

In order to reduce the clock system power, it is effective to reduce a clock voltage swing. Such idea is embodied in the Reduced Clock Swing Flip-Flop (RCSFF) [12]. Figure 3 shows circuit diagrams of the RCSFF. The RCSFF is composed of a current-latch sense amplifier and cross-coupled NAND gates which act as a slave latch. This type of flip-flop was first introduced in 1994 [14] and extensively used in a microprocessor design [13]. The sense-amplifying F/F is often used with low-swing circuits because there is no DC leakage path even if the input is not full swing being different from the conventional gates or F/Fs.

The salient feature of the RCSFF is to accept a reduced voltage swing clock. The voltage swing, Veck, can be as low as 1V. When a clock driver Type A in Fig. 4 is used, power improvement is proportional to Veck, while it is Vclk if Type B driver is used. Type A is easy to implement but less efficient. Type B needs either an internal Veck supply or a DC-DC converter.

The issue of the RCSFF is that when a clock is high to Veck, P1 and P2 do not switch off completely, leaving leak current flowing through either P1 or P2. The power consumption by this leak current turns out to be permissible for some cases (see next section), but further power improvement is possible by reducing the leak current. One way is to apply backgate bias to P1 and P2 and increase the threshold voltage. The other way is to increase the Vth of P1 and P2 by ion-implant, which needs process modification and is usually prohibitive. When the clock is to be stopped, it should be stopped at VSS. Then there is no leak current.

A. Area & Speed

The area of the RCSFF is about 20% smaller than the conventional F/F as seen from Fig. 5 even when the well for the precharge PMOS is separated.

As for delay, SPICE analysis is carried out assuming typical parameters of a generic 0.5µm double metal CMOS process. The delay depends on Wclk (Wclk is defined in Fig.3). Since delay improvement is saturated at Wclk = 10µm, this value of Wclk is used in the area and power estimation. Clock-to-Q delay is improved by a factor of 20% over the conventional F/F even when Vclk = 2.2V, which can be easily realized by a clock driver of the Type A1. Data setup time and hold time in reference to clock are 0.4ns and 0ns, respectively being independent from Vclk, compared to 0.1ns and 0ns for the conventional F/F.

B. Power

The power in the Fig.6 includes clock system power per F/F and the power of a F/F itself. The power consumption is reduced to about 1/2 to 1/3 compared to the conventional F/F depending on the type of the clock driver and V_wl.wl. In the best case studied here, 63% power reduction is observed. TABLE 2 summarizes typical performance improvement.

C. Application to reduced swing bus

For the RCSFF, the D and D̄ input can also be small voltage swing signals. Using this characteristics, the RCSFF can be used to speed up RC delay of long buses. By placing the RCSFF at the end of a long bus and by sense-amplifying the slowly changing D input, RC delay can be reduced to 1/3 compared to the conventional F/F case (see Fig.7).

Let us consider what amount of power gain is observed when a distributed RC line is driven in full swing [15] at one end and switched off when the other terminal becomes VΘ.

\[
V(x,t) = 1 + \frac{x}{2} \sum_{k=1}^{\infty} \left( \frac{1}{k} \right)^{k} \frac{1}{k} \frac{\Gamma \left( k - \frac{1}{2} \right)}{\sqrt{k - \frac{1}{2}}} e^{-\left( k - \frac{1}{2} \right)^{2} x^{2}} \frac{1}{\sqrt{\pi}}
\]

\[
Q = \int_{0}^{L} C \left( V(x,t) dx \right) = CV_{dd} \left[ 1 - \frac{2}{x} \sum_{k=1}^{\infty} \frac{1}{k} \frac{\Gamma \left( k - \frac{1}{2} \right)}{\sqrt{k - \frac{1}{2}}} e^{-\left( k - \frac{1}{2} \right)^{2} x^{2}} \frac{1}{\sqrt{\pi}} \right]
\]

If the energy per cycle, E = (QVdd), is expressed in terms of the terminal voltage, VΘ = (VLq), E = 0.36q + 0.64q. This means that about 50% power saving is possible if an RC interconnect is driven when the voltage swing of VΘ is 0.2Vdd.

4. Low-swing circuit for logic

A pass transistor logic is known to provide a low-power design style. An attempt has been made to further reduce the power-delay product by reducing the signal voltage swing. A Sense-Amplifying Pass-transistor Logic (SAPL) [14] is such a circuit. In the SAPL, a reduced output signal of NMOS pass-transistor logic is amplified by a current latch sense-amplifier to gain speed and save power dissipation as shown in Fig. 9 and Fig.10. The SAPL has been applied to a 1.5ns 20bit carry skip adder in a Discrete Cosine Transform (DCT) macro whose circuit diagram is shown in Fig. 11. 50% speed, 30% area, and 50% power advantage were observed compared with the conventional static CMOS design.

The SAPL is also applied to a 0.9ns 64bit to 32bit double barrel shifter. In this case, 100% speed, 50% area, and 50% power advantage were observed. The MPEG2 decoder LSI which utilizes the DCT and VLD macro with SAPL operates under 0.9V supply voltage.

5. Low-swing circuit for I/O

Application of low-swing circuit to I/O's is also possible. The circuit diagram is shown in Fig.14. The transmitted signal is differential and again is received by a current-latch type sense-amplifier F/F. The two chips are put side by side and bonded directly with minimum capacitance and inductance. The photos of the system are shown in Figs.15 and 16.

At the frequency of 500MHz, the power consumption is 13mW per bonding which includes output and input power (see Fig. 17).
Power: \[ P = p_l \cdot f_{CLK} \cdot C_L \cdot V_{DD}^2 + I_0 \cdot 10^{-10} \cdot \frac{V_{th}}{S} \cdot V_{DD} \]

Delay: \[ \frac{k \cdot Q}{i} = \frac{k \cdot C_L \cdot V_{DD}}{(V_{DD} - V_{th})^\alpha} \]

(a) ![Power vs VDD and VTH](image1)

(b) ![Delay vs VDD and VTH](image2)

Fig. 1 Dependence of (a) power and (b) delay on the supply voltage, \( V_{DD} \) and the threshold voltage, \( V_{TH} \).

### TABLE 1 Comparison of various \( V_{TH} \) control techniques

<table>
<thead>
<tr>
<th>Scheme</th>
<th>MTCMOS</th>
<th>VTCMOS</th>
<th>EVTCMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>St'by</td>
<td>High-Vth</td>
<td>n-well</td>
<td>St'by</td>
</tr>
<tr>
<td></td>
<td>V_{DD}</td>
<td></td>
<td>control</td>
</tr>
<tr>
<td></td>
<td>Low-Vth</td>
<td></td>
<td>(SSB)</td>
</tr>
<tr>
<td></td>
<td>V_{SS}</td>
<td></td>
<td>p-well</td>
</tr>
<tr>
<td>Ref.[3,4]</td>
<td></td>
<td>Ref.[5-9]</td>
<td>Ref.[11]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect</th>
<th>MTCMOS</th>
<th>VTCMOS</th>
<th>EVTCMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ ( I_{St'by} ) reduction</td>
<td>+ ( \Delta V_{th} ) compensation</td>
<td>+ ( \Delta V_{th} ) compensation</td>
<td></td>
</tr>
<tr>
<td>+ ( I_{St'by} ) reduction</td>
<td>+ ( I_{DDQ} ) test</td>
<td>+ ( I_{DDQ} ) test</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penalty</th>
<th>MTCMOS</th>
<th>VTCMOS</th>
<th>EVTCMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>large serial MOSFET(*)</td>
<td>triple well (desirable)</td>
<td>large serial MOSFET operating near threshold(*)</td>
<td></td>
</tr>
<tr>
<td>slower, larger, lower yield</td>
<td>special latch</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 2 Power distribution in VLSI's. MPU1 is a low-end microprocessor for embedded use, MPU2 is a high-end CPU with large amount of cache, ASSP1 is a MPEG2 decoder and ASSP2 is an ATM switch.
(a) RCSFF. Voltage swing of CLK is reduced to Vclk

(b) Conventional F/F

Fig. 3 Circuit diagram of (a) the Reduced Clock Swing Flip-Flop (RCSFF) and (b) the conventional F/F. Numbers in the figure signify MOSFET gate width. Wclk is the gate width of N1.

Table 2 Performance comparison of RCSFF and Conventional F/F

<table>
<thead>
<tr>
<th></th>
<th>Driver</th>
<th>Vclk [V]</th>
<th>Power</th>
<th>Delay</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td></td>
<td>3.3</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>RCSFF</td>
<td>Type A1</td>
<td>2.2</td>
<td>59%</td>
<td>82%</td>
<td>83%</td>
</tr>
<tr>
<td>Wclk = 10μm</td>
<td>Type A2</td>
<td>1.3</td>
<td>48%</td>
<td>123%</td>
<td>83%</td>
</tr>
<tr>
<td>f clk = 100MHz</td>
<td>Type B</td>
<td>2.2</td>
<td>48%</td>
<td>82%</td>
<td>83%</td>
</tr>
</tbody>
</table>

(a) RCSFF (N-well for P1 & P2 separated)

(b) Conventional F/F

Fig. 5 Layout of (a) the Reduced Clock Swing Flip-Flop (RCSFF) with Wclk being 10μm and (b) the conventional F/F.

Fig. 6. Power consumption for one F/F. Clock interconnection length per one F/F is assumed to be 200μm and data activation ratio is assumed to be 30%. fclk is 100MHz. By applying 6V well bias, the initial Vth of P1 and P2 (0.6V) increases to 1.4V.

Fig. 7 Delay improvement of a long RC bus by RCSFF. Wclk = 10μm and Type A1 clock driver is used. Bus is differential and precharged to VDD first and then CLK is asserted when the voltage difference of D and D becomes ΔVD.
Fig. 8 Energy consumed by RC interconnect if the voltage swing of V2 is reduced.

Fig. 10 Timing chart of SAPL

Fig. 12 Waveforms for SAPL adder of Fig. 11

Fig. 9 Sense-Amplifying Pass-Transistor (SAPL) logic concept. The reduced swing signal is amplified by sense-amplifying flip-flop.

Fig. 11 Sense-Amplifying Pass-Transistor (SAPL) applied to 20bit skip-carry adder. The adder was used in a Discrete Cosine Transform macro in a MPEG2 decoder chip which worked under 0.9V.

Fig. 13 Sense-Amplifying Pass-Transistor (SAPL) applied to a 32bit barrel shifter. The shifter was used in a Variable Length Decoder macro in a MPEG2 decoder chip which worked under 0.9V.
Fig. 14  Circuit diagram of low-swing I/O. The upper half is a transmitter side and the lower half is a receiver side.

Fig. 15  Microphotograph showing bonding pads and I/O circuits (mostly under Al lines). The I/O circuit includes input and output circuit and is smaller than a pad.

Fig. 16  Photograph to show two chips are connected by bonding wires directly.

Fig. 17  Measured waveform on the bonding pads. The frequency is 500MHz.

References