22.8 Zigzag Super Cut-off CMOS (ZSCCMOS) Block Activation with Self-Adaptive Voltage Level Controller: An Alternative to Clock-Gating Scheme in Leakage Dominant Era

Kyeong-Sik Min[1], Hiroshi Kawaguchi, Takayasu Sakurai

University of Tokyo, Tokyo, Japan
[1]Now wih Kookmin University, Seoul, Republic of Korea

Leakage power is expected to be dominant in sub-70nm technology. In the leakage dominant environments, a clock gating scheme loses its effectiveness although currently it is the most effective way to reduce power consumption. A clock-gating substitute is proposed which achieves a 200ps wake-up time and 3 orders of magnitude leakage reduction for leakage dominant LSI's and experimentally verifies the effectiveness of the scheme.

Super Cut-off CMOS (SCCMOS) in Fig. 22.8.1a, which does not use a high-V_{TH} switch, can be operated at sub-1V V_{DD} [1] and is scalable. However, it suffers from a long wake-up time and a high current peak at the sleep-to-active transition. This prevents the scheme from being used as a clock-gating substitute with less than one cycle latency. Node (L) voltage and V_{SSV} (virtual V_{SS} line) are initially low but they become high during sleep because the switch, MN, is cut off completely. Once they become high, the charge associated with all low internal gate nodes and V_{SSV} is restored low at wake-up. This makes the wake-up time long and the peak current high. Figure 22.8.1b shows Zigzag Super Cut-off CMOS (ZSCCMOS) that is proposed to improve the operating speed by eliminating the series-connected switches while achieving the relaxation of the high-voltage stress at the cut-off switch [2]. The zigzag configuration reduces the wake-up time and rush current, because the voltage of node (L) does not change during sleep and V_{SSV} changes very little by means of a self-reverse-biasing effect. Though this zigzagging concept is previously proposed in [3], the cut-off switch with high V_{TH} degrades both the wake-up time and operating speed severely and moreover, the wake-up feature was not discussed in [3].

The same effects of ZSCCMOS are expected for another zigzag scheme called Zigzag Boosted Gate MOS (ZBGMOS) and shown in Fig. 22.8.1c. The MN and MP switch MOSFET's are made with thick gate oxide and high V_{TH}. Unlike the ZSCCMOS, the power consumption in generating V_{GP} and V_{GN} which are higher and lower than V_{DD} and V_{SS}, respectively, is zero in a sleep mode but finite in an active mode (although it is negligible in most applications). Figure 22.8.1d shows a voltage stress plot of the cut-off switch, MN, in Figs. 22.8.1a and 22.8.1b. This shows that a voltage over-stress higher than V_{DD} does not occur as long as $|V_{GN}|$ is smaller than 0.3V in the ZSCCMOS. Alternatively, in the conventional SCCMOS, $|V_{GD}|$ begins to exceed V_{DD} when $|V_{GN}|$ is more than 0.05V and this small allowance of $|V_{GN}|$ limits the cut-off magnitude.

In this clock-gating method both the switching and leakage power are reduced by block-wise activation. The ZSCCMOS block activation scheme is composed of a combinational logic block (CLB), F/F's, and a control circuit as shown in Fig. 22.8.2a. A local clock (LCLK) is controlled by a block enable signal (EN). The fast recovery feature of the ZSCCMOS is best suited to block activation where the fast wake-up time of less than a cycle is essential. Figure 22.8.2b shows the schematic waveforms for Fig.22.8.2a. V_{SSV} should be restored to V_{SS} before the LCLK arrives at the input F/F, implying the necessity of the fast wake-up. To design the combinational logic block using ZSCCMOS, every node voltage in the block should be predictable at the sleep mode, since cut-off switches should be inserted in series to the off MOSFET's in the combinational block. Unfortunately, node

voltages are generally not predictable, since they are determined by the last inputs to the block. To make every node voltage predictable, a new phase-forcing circuit is added to the input F/F depicted in Fig. 22.8.3a, where P and Q are forced to be low or high, even though D is unknown during sleep. Without input phase forcing, the internal node voltages of the block may be changing from H to L or L to H during sleep. When the wake-up signal comes, if the internal node voltages are in between V_{DD} and V_{SS}, a large short-circuit peak current may arise and large wake-up delay ruins the scheme.

Figure 22.8.3b shows a leakage-suppressed output F/F. Since Q of the output F/F is connected to the other logic block driven by a different LCLK or a global clock (GCLK), the Q is not be forced to H or L and keeps its state during the sleep. Therefore, G1, G2, and G3 are not be connected to the virtual power lines. Since G2 and G3 are not on the critical path, they can be built with high-V_{TH} devices. G1 is connected to both NMOS and PMOS cut-off switches to suppress the leakage.

Figure 22.8.4 shows a V_N (negative voltage) generator for the NMOS cut-off switch. It is composed of a self-adaptive voltage-level detector (SAVLD) and a charge pump. In addition, an oxide-stress-relaxed level shifter is also shown in Fig. 22.8.4, where negative voltage lower than V_{SS} is distributed to V_{GN} without giving rise to any over-stress [4]. In the SAVLD, if V_N is too close to V_{SS}, VLD_O turns on the charge pump to lower V_N. It should be noted that V_N changes self-adaptively with the process, voltage and temperature variation. The lower the V_{GN} of the cut-off switch is, the smaller the leakage is but at the same time, the lower V_{GN} pushes up V_{SSV} higher during sleep and thus increases the wake-up time and rush current. The self-adaptive V_N generator makes a trade-off between wake-up time and leakage suppression. Figure 22.8.5a shows the measured waveforms of the V_P generator fabricated in the 0.6μm technology when V_{DD} =1.1V. If V_P goes lower than a target voltage of 1.2V even by 30mV, the SAVLD turns on the charge pump. Figure 22.8.5b shows the measured ΔV_P with varying the temperature from 25°C to 75°C. This figure shows good agreement between measurement and simulation.

Figure 22.8.6a shows the measured wake-up time of the ZSCCMOS and SCCMOS that are fabricated using the 0.6μm technology. The comparison shows that the wake-up time of the ZSCCMOS is 8 times faster than that of the SCCMOS. The inverter delay with a fan-out of 3 is also shown in Fig. 22.8.6a. From this figure, it is seen that the wake-up time of ZSCCMOS is almost the same with the inverter delay with a fan out of 3. Figure 22.8.6b compares the V_{SSV} wake-up time of various schemes in future designs, where the V_{SSV} wake-up time is defined by the time of V_{SSV} being restored to 95% of V_{SS}. Simulation is carried out by using the projected 70nm MOS models provided by the BPTM [5]. The wake-up time of the ZSCCMOS is 12 times faster than the SCCMOS in 70nm technology. As in Fig. 22.8.6a, the wake-up time of ZSCCMOS in 70nm technology in Fig. 22.8.6b seems to be almost the same with the inverter delay with a fan out of 3.

References
[1] H. Kawaguchi et al., *IEEE J. Solid State Circuits*, vol. 35, pp. 1498-1501, Oct. 2000.
[2] K. Min et al., *28th European Solid-State Circuits Conference*, pp. 679-682, Sep. 2002.
[3] M. Horiguchi et al., *IEEE J. Solid State Circuits*, vol. 28, pp. 1131-1135, Nov. 1993.
[4] Y. Nakagome et al., *IEEE J. Solid State Circuits*, vol. 26, pp. 1003-1010, July 1991.
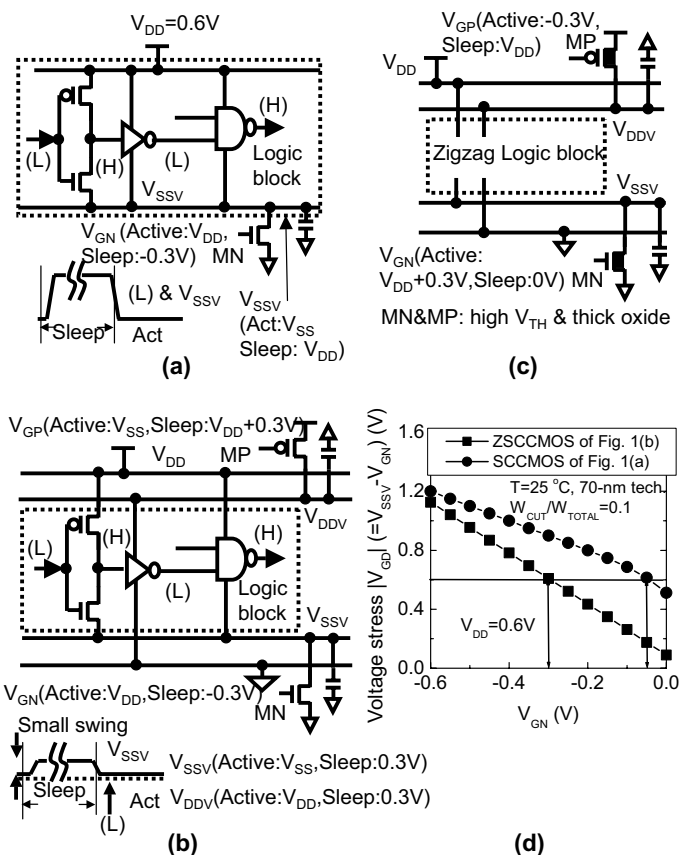[5] Berkeley Predictive Technology Model web site (BPTM)

Figure 22.8.1: (a) Super cut-off CMOS (SCCMOS) Scheme (b) Zigzag super cut-off CMOS (ZSCCMOS) Scheme (c) Zigzag Boosted Gate MOS (ZBGMOS) Scheme (d) Voltage stress plot of the cut-off switch of Figure 22.8.1(a) and (b). WCUT and WTOTAL mean the width of cut-off switch and the total width of MOSFET's that are connected to the cut-off switch, respectively.

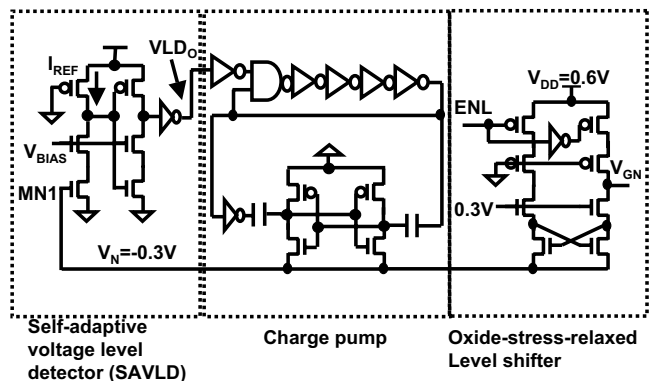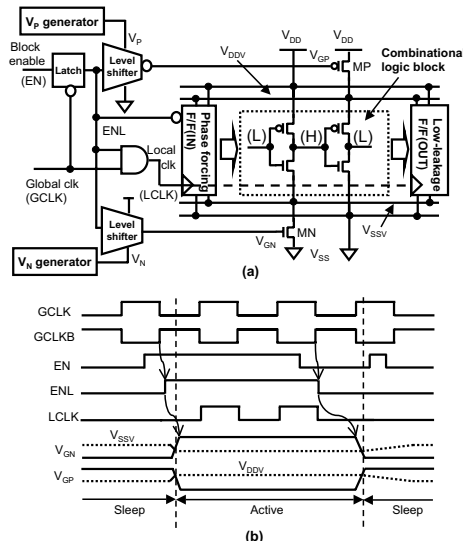Figure 22.8.2: (a) ZSCCMOS block activation scheme (b) Its waveforms.
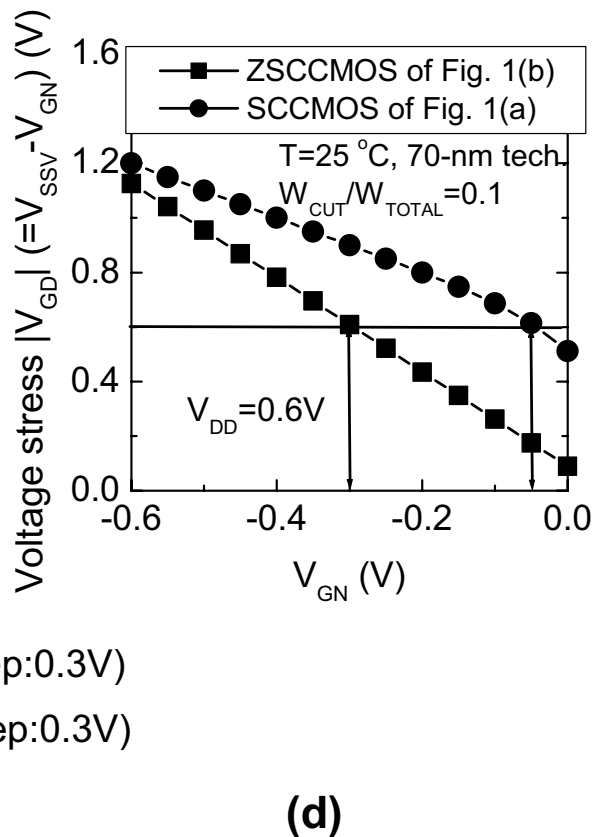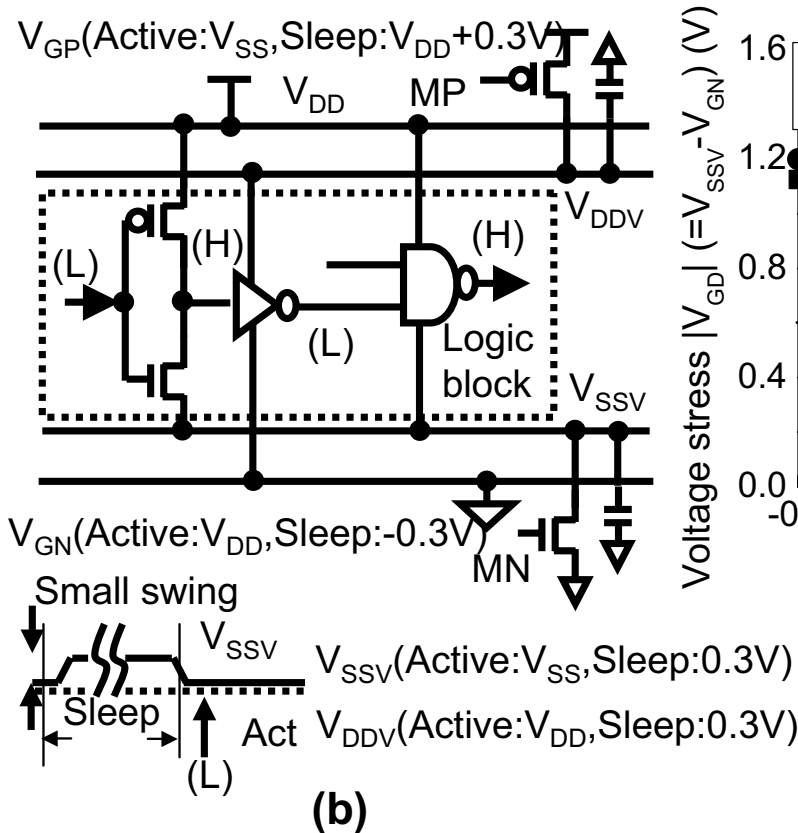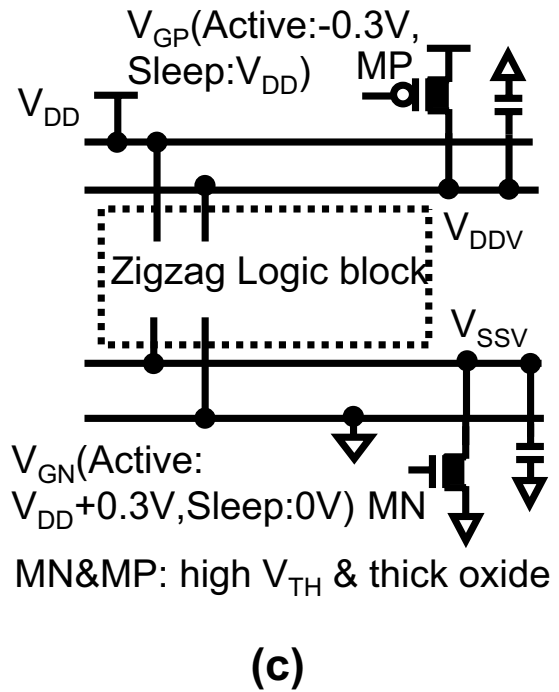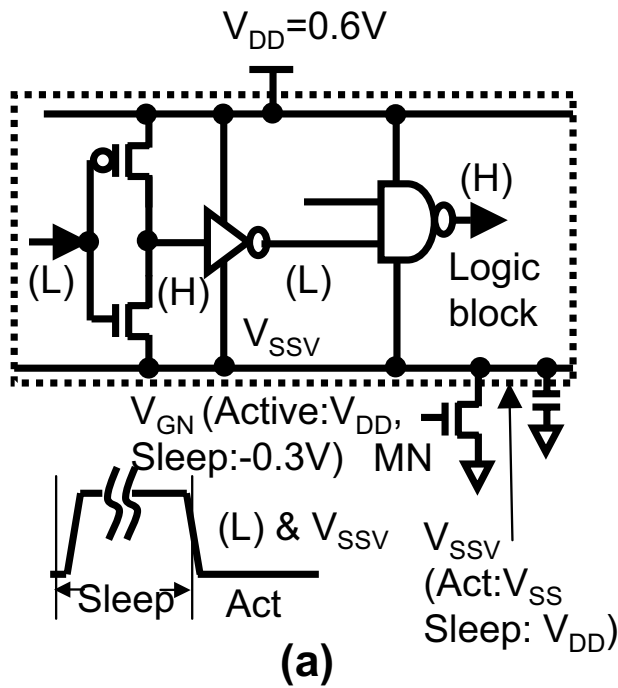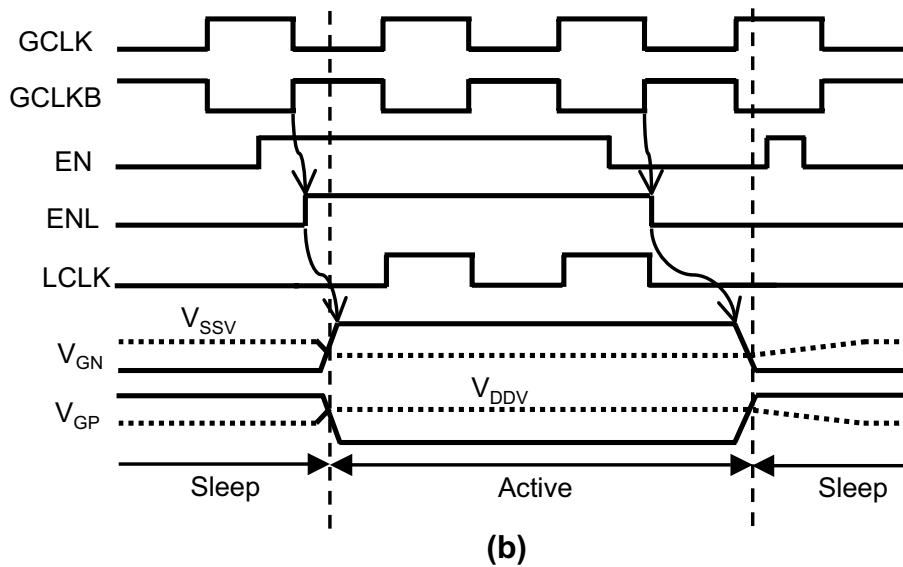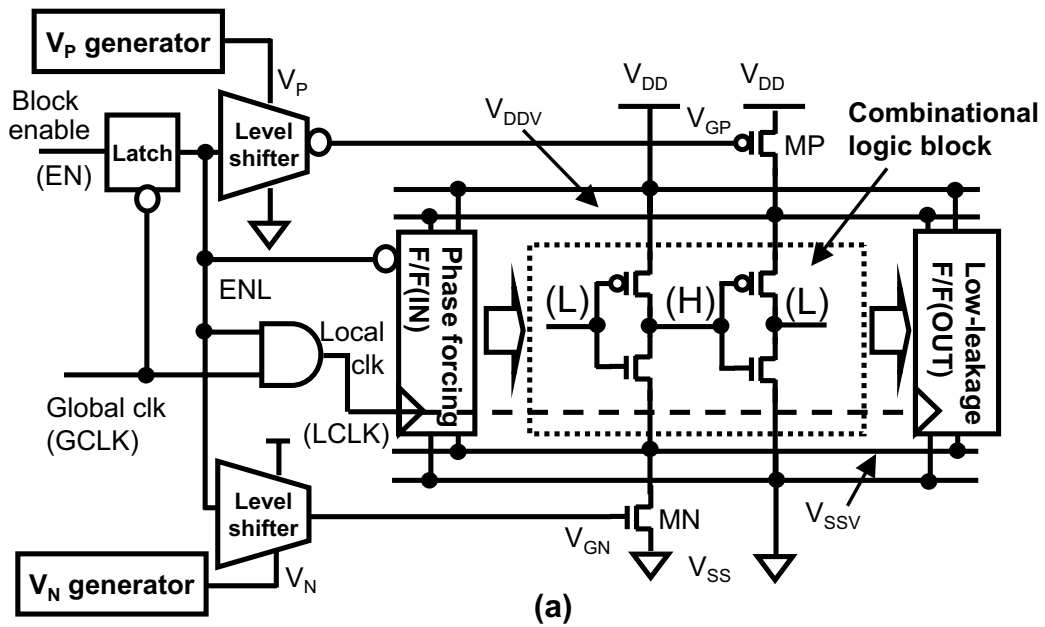

Figure 22.8.3: (a) Input flip-flop with phase-forcing circuit (b) Low-leakage output flip-flop.

Figure 22.8.4: V_N (negative voltage) generator for NMOS cut-off switch with self-adaptive voltage level detector, charge pump, and oxide-stress-relaxed level shifter.

Figure 22.8.5: (a) Measured operating waveforms of V_P (positive voltage) generator (b) Measured ΔV_P with varying temperature.

22
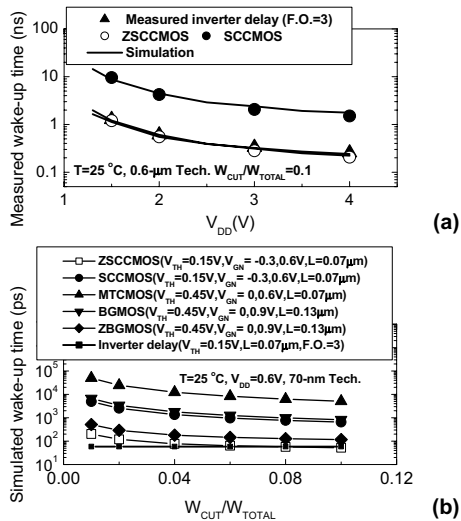
Figure 22.8.6: (a) Measured wake-up times in 0.6μm technology; (b) Simulated wake-up times in the projected 70-nm technology.

Figure 22.8.7: Test chip micrograph.

22

Figure 22.8.1: (a) Super cut-off CMOS (SCCMOS) Scheme (b) Zigzag super cut-off CMOS (ZSCCMOS) Scheme (c) Zigzag Boosted Gate MOS (ZBGMOS) Scheme (d) Voltage stress plot of the cut-off switch of Figure 22.8.1(a) and (b). WCUT and WTOTAL mean the width of cut-off switch and the total width of MOSFET's that are connected to the cut-off switch, respectively.

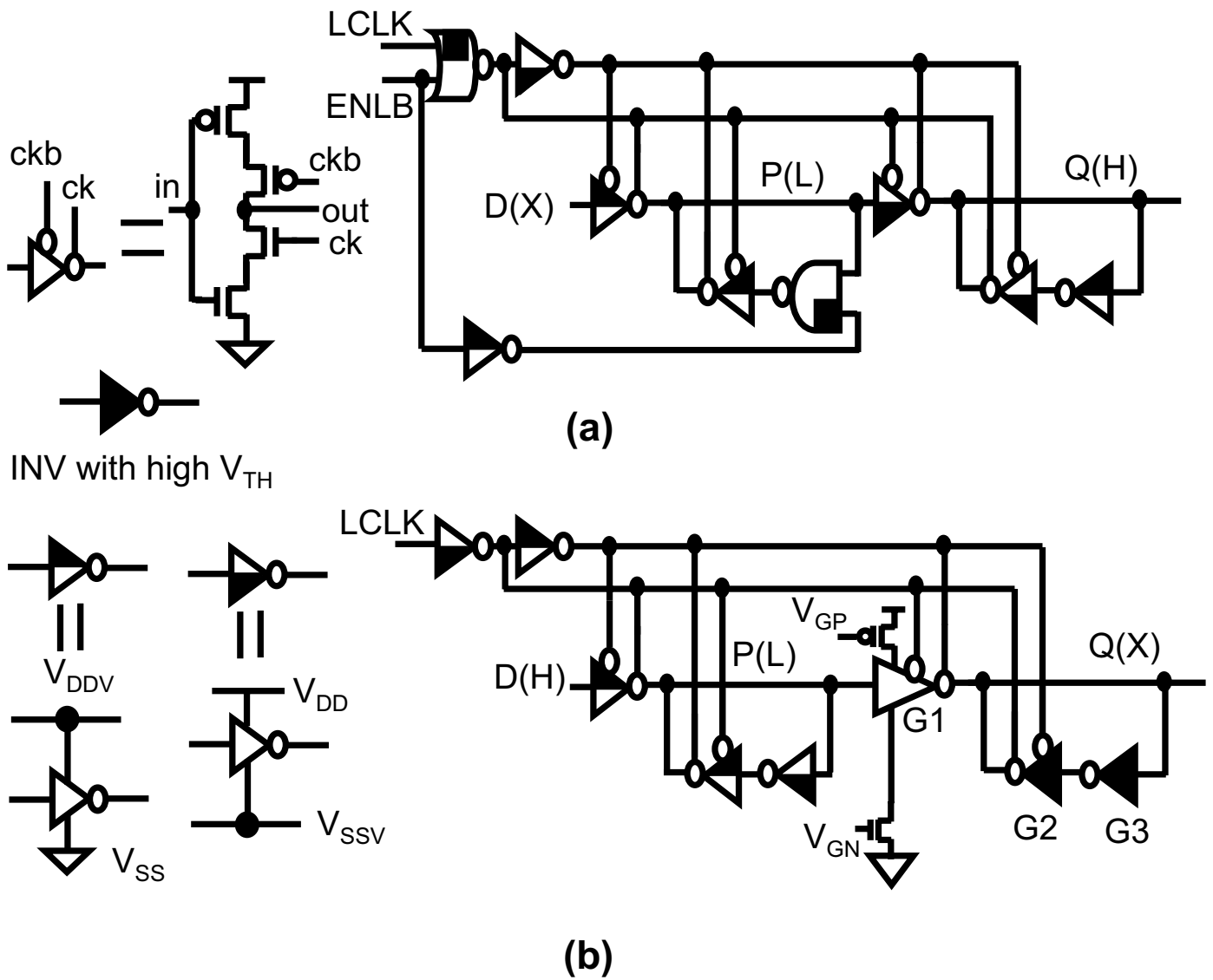Figure 22.8.2: (a) ZSCCMOS block activation scheme (b) Its waveforms.

Figure 22.8.3: (a) Input flip-flop with phase-forcing circuit (b) Low-leakage output flip-flop.
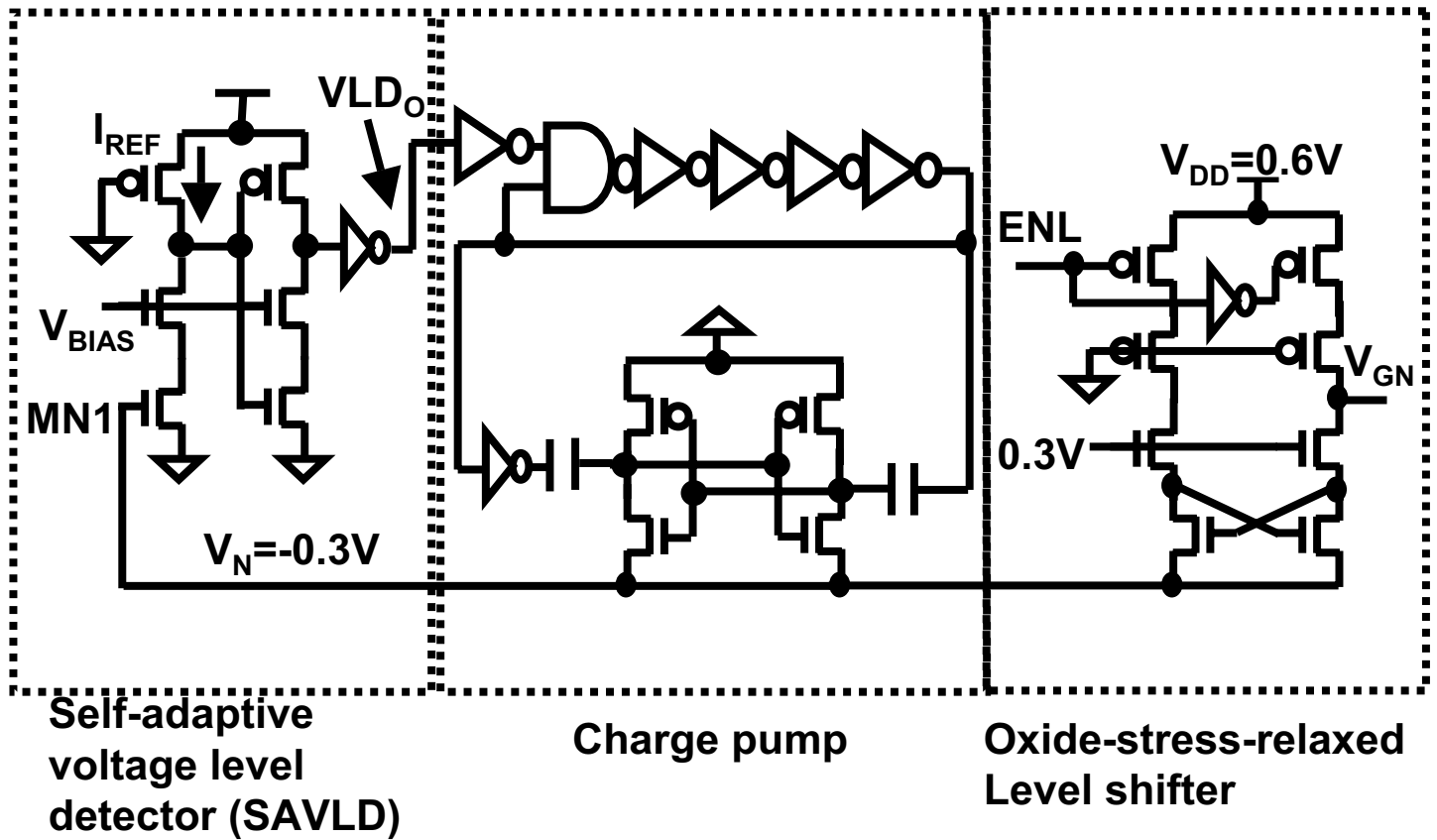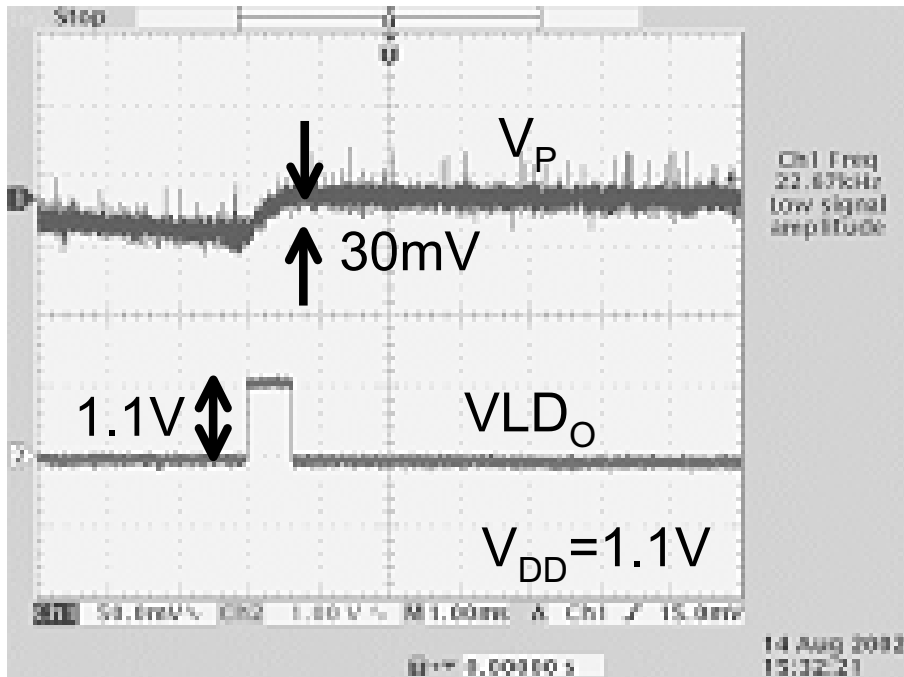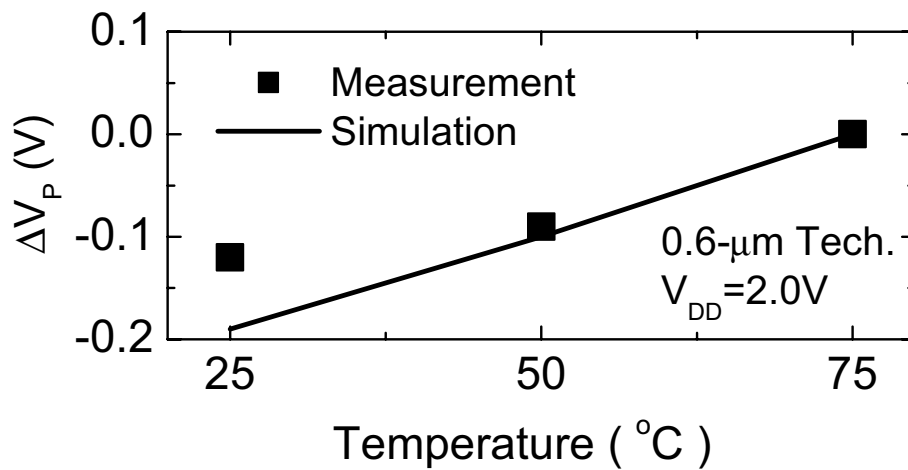
Figure 22.8.4: V_N (negative voltage) generator for NMOS cut-off switch with self-adaptive voltage level detector, charge pump, and oxide-stress-relaxed level shifter.

(a)

(b)

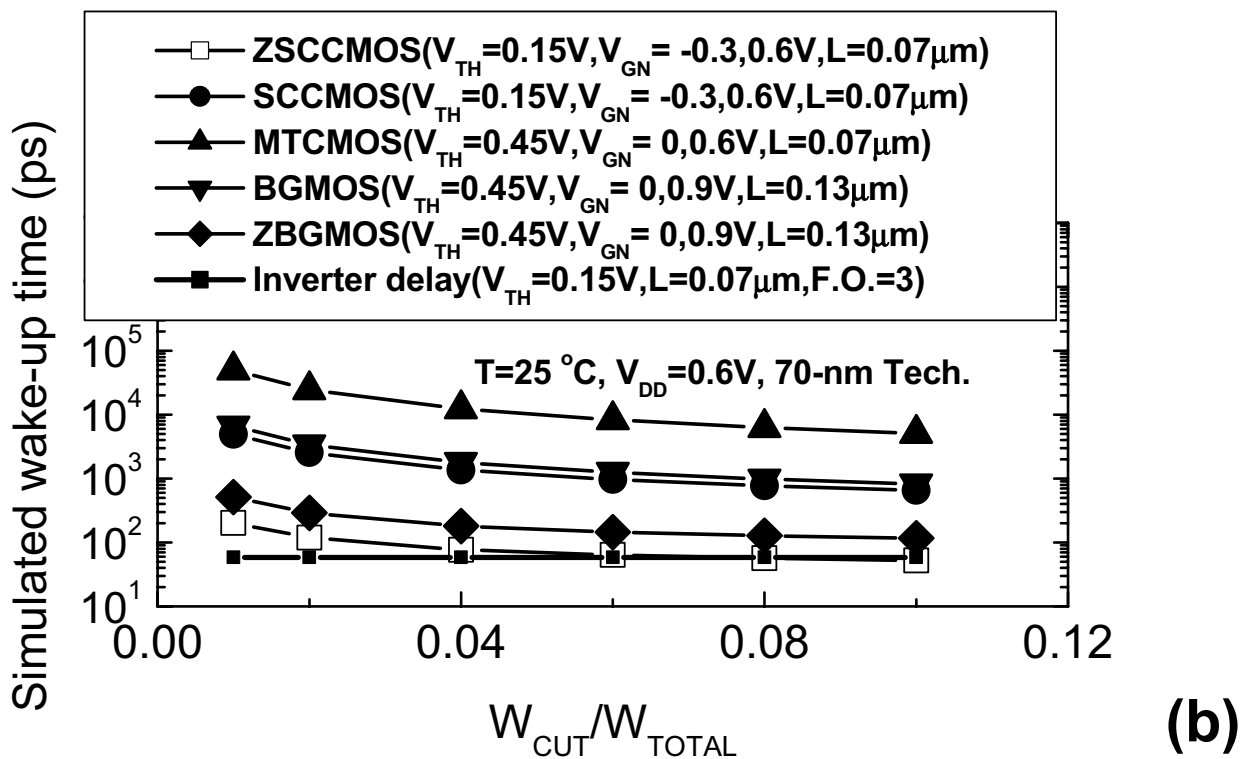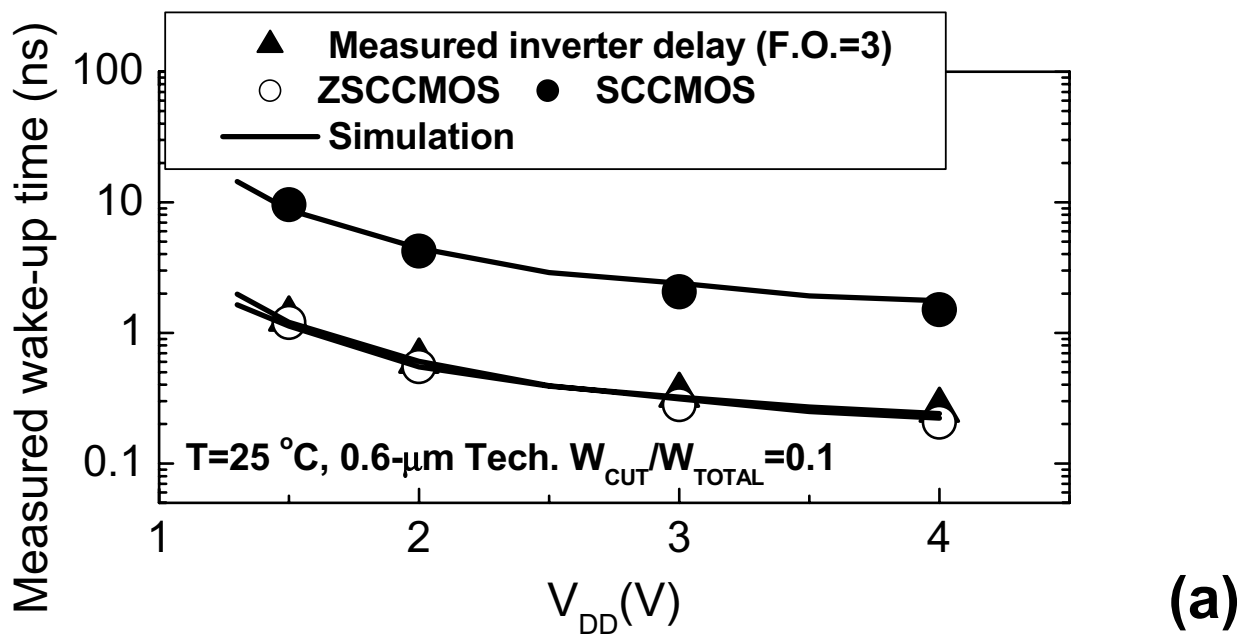Figure 22.8.5: (a) Measured operating waveforms of V_P (positive voltage) generator (b) Measured ΔV_P with varying temperature.

Figure 22.8.6: (a) Measured wake-up times in 0.6μm technology; (b) Simulated wake-up times in the projected 70-nm technology.

I/O buffers

Ring configuration to measure wake-up time
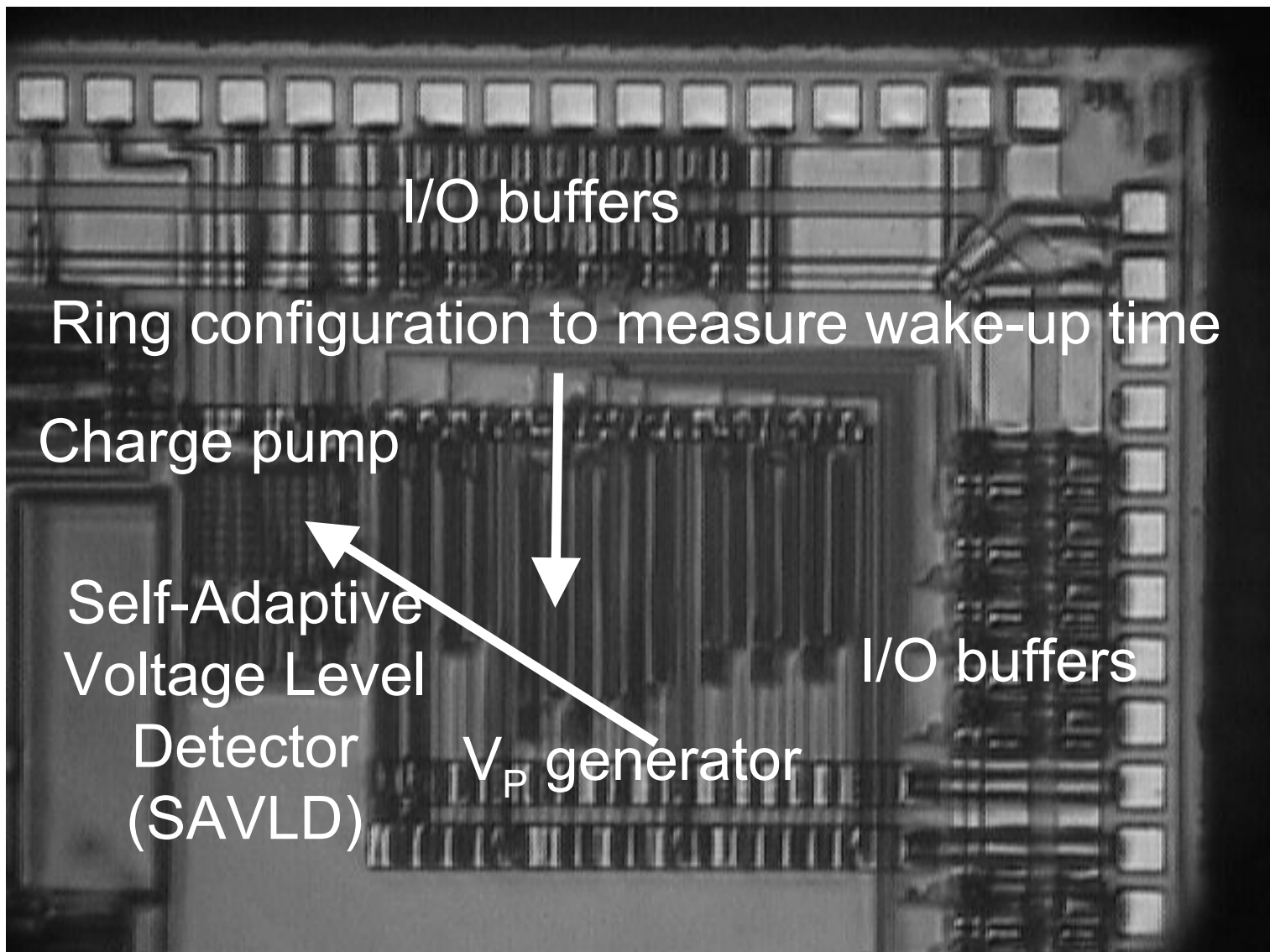
Charge pump

Self-Adaptive Voltage Level Detector (SAVLD)

V_P generator

I/O buffers

Figure 22.8.7: Test chip micrograph.