Clock-Gating and Its Application to Low Power Design of Sequential Circuits

Qing WU
Department of Electrical Engineering-Systems, University of Southern California
Los Angeles, CA 90089, USA, Phone: (213)740-4480

Massoud PEDRAM
Department of Electrical Engineering-Systems, University of Southern California
Los Angeles, CA 90089, USA, Phone: (213)740-4458

Xunwei WU
Department of Electronic Engineering, Hangzhou University
Hangzhou, Zhejiang 310028, CHINA

ABSTRACT

This paper models the clock behavior in a sequential circuit by a quaternary variable and uses this representation to propose and analyze two clock-gating techniques. It then uses the covering relationship between the triggering transition of the clock and the active cycles of various flip-flops to generate a derived clock for each flip-flop in the circuit. Design examples using gated clocks are provided next. Experimental results show that these designs have ideal logic functionality with lower power dissipation compared to traditional designs.

Clock-Gating and Its Application to Low Power Design of Sequential Circuits

I. INTRODUCTION

The sequential circuits in a system are considered major contributors to the power dissipation since one input of sequential circuits is the clock, which is the only signal that switches all the time. In addition, the clock signal tends to be highly loaded. To distribute the clock and control the clock skew, one needs to construct a clock network (often a clock tree) with clock buffers. All of this adds to the capacitance of the clock net. Recent studies indicate that the clock signals in digital computers consume a large (15% - 45%) percentage of the system power (1). Thus, the circuit power can be greatly reduced by reducing the clock power dissipation.

Most efforts for clock power reduction have focused on issues such as reduced voltage swings, buffer insertion and clock routing (2). In many cases switching of the clock causes a lot of unnecessary gate activity. For that reason, circuits are being developed with controllable clocks. This means that from the master clock other clocks are derived which, based on certain conditions, can be slowed down or stopped completely with respect to the master clock. Obviously, this scheme results in power savings due to the following factors:

1) Load on the master clock is reduced and the number of required buffers in the clock tree is decreased. Therefore, the power dissipation of clock tree can be reduced.

2) The flip-flop receiving the derived clock is not triggered in idle cycles; the corresponding dynamic power dissipation is thus saved.

3) The excitation function of the flip-flop triggered by derived clock may be simplified since it has a don’t care condition in the cycle when the flip-flop is not triggered by the derived clock.

In (3) the authors presented a technique for saving power in the clock tree by stopping the clock fed into idle modules. However, a number of engineering issues related to the design of the clock tree were not addressed and hence, the proposed approach has not been adopted in practice.

This paper investigates various issues in deriving a gated clock from a master clock. In section II, a quaternary variable is used to model the clock behavior and to discuss its triggering action on flip-flops. Based on this analysis, two clock-gating schemes are proposed. In section III, we use the covering relation between the clock and the transition behaviors of the triggered flip-flops to derive conditions for gating the master clock. Two common sequential circuits, i.e. 8421 BCD code up-counter and three-excess counter, are then described to illustrate the procedure for finding a derived clock. In section IV, a new technique for clock-gating is presented

which generates a clock synchronous with the master clock. This eliminates the additional skew between the master clock and the derived clock. Thus, the designed sequential circuit is a synchronous one. Finally, we present circuit simulation results to prove the quality of the derived clock and its ability to reduce power dissipation in the circuit.

II. DESCRIPTION FOR CLOCK BEHAVIOR AND CLOCK-GATING

In a synchronous system, a flip-flop is triggered by a certain directional transition of a clock signal. For the clock to be another signal rather than the master clock, it must offer the same directional transition to trigger the flip-flop, and it must be "in step" with the master clock.

For the clock signal *clk* in a circuit if we denote its logic values before and after a transition as $clk(t)$ and $clk^+(t)$ respectively, four combinations can be used to express different behaviors of the clock as shown in Table 1, where a special quaternary variable $\tilde{clk}$ denotes the corresponding behavior. The four values are (0, α, β, 1), where α, β represent two kinds of transition behaviors and 0, 1 represent two kinds of holding behaviors. (Note that although they have the same forms as signal values 0 and 1, their meanings are different.)

Table 1 QUATERNARY REPRESENTATION FOR BEHAVIORS OF A SIGNAL

$\tilde{clk}$	$clk(t) \rightarrow clk^+(t)$		Behavior
0	0	0	0-holding
α	0	1	α-transition
β	1	0	β-transition
1	1	1	1-holding

In addition, we can also define a *literal* operation to identify the behavior of a clock:

$$clk^b = \begin{cases} 1 & \text{if } \tilde{\text{clk}} = b \\ 0 & \text{if } \tilde{\text{clk}} \neq b, \end{cases} \tag{1}$$

where $b \in \{0, \alpha, \beta, 1\}$. Thus, the rising transition clk^α and the falling transition clk^β of a clock are binary variables and can serve as arguments of Boolean operations. For example, from Table 1 we have

$clk^0 = \overline{clk} \cdot \overline{clk^+}, clk^\alpha = \overline{clk} \cdot clk^+, clk^\beta = clk \cdot \overline{clk^+}$ and $clk^1 = clk \cdot clk^+$.

Assume that there are n flip-flops in a sequential circuit and that their outputs and clock inputs are denoted by Q_i and clk_i, i = 0,1,…,n-1, respectively. For a synchronous sequential circuit, we have $clk_i = clk$, namely all flip-flops are triggered by the same master clock signal clk. However, if a flip-flop Q_i is to be disconnected from the master clock during some (idle) cycles, then we have to use a derived clock for Q_i. Notice that this derived clock should be "in step" with the master clock for the circuits to remain synchronous.

Generally, we consider that the derived clock is obtained from the master clock clk and the outputs of other flip-flops $Q_0,\cdots,Q_{i-1},Q_{i+1},\cdots,Q_{n-1}$, (which make transitions following the triggering transition of their respective clocks.) Since both AND gating and OR gating can be used for controlling the master clock, we have the following two clock-gating forms

$$clk_i = g_i + p_i \cdot clk\,, \tag{2}$$

$$clk_i = g_i \cdot (p_i + clk)\,, \tag{3}$$

where g_i and p_i are functions of flip-flop outputs $Q_0,\cdots,Q_{i-1},Q_{i+1},\cdots,Q_{n-1}$.

Consider a flip-flop triggered by the falling clock transition as an example (i.e. a negative edge-triggered flip-flop). The timing relationships of clk, p_i, $p_i \cdot clk$ and $p_i + clk$ are shown in Fig.1. Note that p_i exhibits a delay with respect to the falling transition of clock, may have glitches (represented by vertical grid lines), and has its final stable value in the zone where clk = 0. We can see that $p_i + clk$ cannot prevent the glitches and may even lead to an extra glitch. Therefore, only (2) is suitable for the negative edge-triggered flip-flops while (3) is not. Note that g_i in (2) must be glitch-free when clk = 0.

The above discussion shows that the falling transition of clk_i in (2) occurs for the following two cases:

(1) When g_i = 0 and p_i = 1, falling transition of clk leads to falling transition of the derived clock clk_i. Therefore, p_i may be named the *transition propagate* term.

(2) When g_i = 1 and g_i makes a falling transition, the derived clock clk_i makes a falling transition since clk and hence $p_i \cdot clk$ are 0 at that time instance. Therefore, g_i may be named the *transition generate* term.

Figure 1 Timing relationship of *clk*, p_i(g_i), p_i.clk and p_i+clk

From this analysis, we obtain

$$clk_i^{\beta} = g_i^{\beta} + \overline{g_i} \cdot p_i \cdot clk^{\beta}. \quad (4)$$

Similarly, we can show that the derived clock signal in (3) is suitable for the flip-flops triggered by the rising transition of the clock. Here g_i in (3) must be glitch-free when clk = 1. The rising transition of clk_i can be expressed as

$$clk_i^{\alpha} = g_i^{\alpha} + g_i \cdot \overline{p_i} \cdot clk^{\alpha}. \quad (5)$$

It should be pointed out that the attached circuitry needed for generating the derived clock should be simple to avoid excessive power dissipation due to this overhead circuitry. Therefore g_i and p_i in (2) and (3) should be relatively simple functions. Especially, we require g_i to be simple to avoid dangerous glitches. Note that if g_i = 0, p_i = 1 in (2) or g_i = 1, p_i = 0 in (3), we return to the condition of applying the master clock clk in a synchronous sequential circuit.

III. DESIGN OF SEQUENTIAL CIRCUITS BASED ON DERIVED CLOCK

Assume that the derived clock for the flip-flop Q_i is clk_i. Falling transitions of clk_i have to cover all cycles when the flip-flop Q_i makes transitions, Q_i^{α} and Q_i^{β}. The covering relation can be expressed as:

$$clk_i^{\beta} \geq (Q_i^{\alpha} + Q_i^{\beta}). \quad (6)$$

Since AND and OR operations on Boolean variables can be interpreted as minimum and maximum operations on these variables, i.e. $x \cdot y = \min(x, y)$ and $x + y = \max(x, y)$, we can obtain the following equations from (6)

$$clk_i^{\beta} \cdot (Q_i^{\alpha} + Q_i^{\beta}) = (Q_i^{\alpha} + Q_i^{\beta}), \quad (7)$$

$$clk_i^{\beta} + (Q_i^{\alpha} + Q_i^{\beta}) = clk_i^{\beta}. \quad (8)$$

Therefore, we should first obtain $(Q_i^{\alpha} + Q_i^{\beta})$ and then generate the derived clock clk_i for flip-flop Q_i. We will show the procedure by using design examples.

Example 1. Design of an 8421 BCD code up-counter

The next states and state behaviors of an 8421 BCD code up-counter are shown in Table 2, where behavior of each flip-flop ($Q_i \rightarrow Q_i^+$) is denoted by $\tilde{Q}_i$. From Table 2, the corresponding next state Karnaugh maps and behavior Karnaugh maps may be obtained, as shown in Fig.2(a) and 2(b). In these maps an empty box represents

the don't care condition. The two transition functions for each flip-flop can be derived from their corresponding behavior Karnaugh maps as below:

Table 2 NEXT STATES AND STATE BEHAVIORS OF AN 8421 BCD CODE UP-COUNTER

Q_3	Q_2	Q_1	Q_0	Q_3^+	Q_2^+	Q_1^+	Q_0^+	$\tilde{Q}_3$	$\tilde{Q}_2$	$\tilde{Q}_1$	$\tilde{Q}_0$
0	0	0	0	0	0	0	1	0	0	0	α
0	0	0	1	0	0	1	0	0	0	α	β
0	0	1	0	0	0	1	1	0	0	1	α
0	0	1	1	0	1	0	0	0	α	β	β
0	1	0	0	0	1	0	1	0	1	0	α
0	1	0	1	0	1	1	0	0	1	α	β
0	1	1	0	0	1	1	1	0	1	1	α
0	1	1	1	1	0	0	0	α	β	β	β
1	0	0	0	1	0	0	1	1	0	0	α
1	0	0	1	0	0	0	0	β	0	0	β

$$Q_3^{\alpha} = Q_2 \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}, Q_3^{\beta} = Q_3 \cdot Q_0 \cdot clk^{\beta}; \tag{9}$$

$$Q_2^{\alpha} = \overline{Q_2} \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}, Q_2^{\beta} = Q_2 \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}; \tag{10}$$

$$Q_1^{\alpha} = \overline{Q_3} \cdot \overline{Q_1} \cdot Q_0 \cdot clk^{\beta}, Q_1^{\beta} = \overline{Q_3} \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}; \tag{11}$$

$$Q_0^{\alpha} = \overline{Q_0} \cdot clk^{\beta}, \; Q_0^{\beta} = Q_0 \cdot clk^{\beta}. \tag{12}$$

Therefore, we have

$$Q_3^{\alpha} + Q_3^{\beta} = (Q_3 + Q_2 \cdot Q_1) \cdot Q_0 \cdot clk^{\beta}, \tag{13}$$

$$Q_2^{\alpha} + Q_2^{\beta} = Q_1 \cdot Q_0 \cdot clk^{\beta}, \tag{14}$$

$$Q_1^{\alpha} + Q_1^{\beta} = \overline{Q_3} \cdot Q_0 \cdot clk^{\beta}, \tag{15}$$

$$Q_0^{\alpha} + Q_0^{\beta} = clk^{\beta}. \tag{16}$$

From (13)-(15) we find that $Q_0 \cdot clk^{\beta} \geq (Q_i^{\alpha} + Q_i^{\beta})$, (i =1,2,3). From (12), we see that $Q_0^{\beta} = Q_0 \cdot clk^{\beta}$ can serve as the needed falling transition trigger for flip-flops Q_1, Q_2, and Q_3, namely $clk_1^{\beta} = clk_2^{\beta} = clk_3^{\beta} = Q_0^{\beta}$. Comparing these with (4), we get $g_i = Q_0$, $p_i = 0$ and $clk_i = Q_0$. (i =1,2,3).

As for clk_0, (16) indicates that the clock for Q_0 is no other than the master clock *clk*. Since we only need take care of the excitation input when the flip-flop receives a triggering falling clock transition (i.e. β entries in $\tilde{Q}_0$ map), we don't care what the excitation inputs in other conditions are. Therefore the next state Karnaugh maps for flip-flops Q_1, Q_2, and Q_3 in Fig.2(a) can be simplified to those shown in Fig.2(c).

Figure 2

(a) Next state Karnaugh maps, (b) behavior Karnaugh maps, (c) simplified next state Karnaugh maps

From Fig.2(a),(c) we can get the corresponding both synchronous and asynchronous designs, as shown in Fig.3. (We say asynchronous, because now not all flip-flops are triggered at the same time.) Obviously the corresponding combinational circuits are simpler. Besides, since three flip-flops Q_3, Q_2, Q_1 have no dynamic power dissipation half of the time when there is no clock triggering, and because the simpler combinational circuits has lower node capacitance, the asynchronous design is saving power.

Figure 3 Circuit realizations of BCD code up-counter

(a) synchronous design, (b) asynchronous design

Example 2. Design of an excess-three code up-counter

The next state and state transition of an excess-three code up-counter are shown in Table 3. Transition functions for each flip-flop can be derived as below

$$Q_3^{\alpha} = Q_2 \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}, \quad Q_3^{\beta} = Q_3 \cdot Q_2 \cdot \overline{Q_1} \cdot clk^{\beta}; \tag{17}$$

$$Q_2^{\alpha} = \overline{Q_2} \cdot Q_1 \cdot Q_0 \cdot clk^{\beta}, \; Q_2^{\beta} = (Q_3 \cdot Q_2 \cdot \overline{Q_1} + Q_2 \cdot Q_1 \cdot Q_0) \cdot clk^{\beta}; \tag{18}$$

$$Q_1^{\alpha} = (Q_3 \cdot Q_2 \cdot \overline{Q_0} + \overline{Q_1} \cdot Q_0) \cdot clk^{\beta}, \; Q_1^{\beta} = Q_1 \cdot Q_0 \cdot clk^{\beta}; \tag{19}$$

$$Q_0^{\alpha} = \overline{Q_0} \cdot clk^{\beta}, \quad Q_0^{\beta} = Q_0 \cdot clk^{\beta}. \tag{20}$$

Table 3 THE NEXT STATES AND STATE BEHAVIORS OF A EXCESS-THREE CODE UP-COUNTER.

Q_3	Q_2	Q_1	Q_0	Q_3^+	Q_2^+	Q_1^+	Q_0^+	$\tilde{Q}_3$	$\tilde{Q}_2$	$\tilde{Q}_1$	$\tilde{Q}_0$
0	0	1	1	0	1	0	0	0	α	β	β
0	1	0	0	0	1	0	1	0	1	0	α
0	1	0	1	0	1	1	0	0	1	α	β
0	1	1	0	0	1	1	1	0	1	1	α
0	1	1	1	1	0	0	0	α	β	β	β
1	0	0	0	1	0	0	1	1	0	0	α
1	0	0	1	1	0	1	0	1	0	α	β
1	0	1	0	1	0	1	1	1	0	1	α
1	0	1	1	1	1	0	0	1	α	β	β
1	1	0	0	0	0	1	1	β	β	α	α

Therefore, we have

$$Q_3^{\alpha} + Q_3^{\beta} = (Q_3 \cdot Q_2 \cdot \overline{Q_1} + Q_2 \cdot Q_1 \cdot Q_0) \cdot clk^{\beta} = Q_2^{\beta}, \tag{21}$$

$$Q_2^{\alpha} + Q_2^{\beta} = (Q_3 \cdot Q_2 \cdot \overline{Q_1} + Q_1 \cdot Q_0) \cdot clk^{\beta} = Q_1^{\beta} + \overline{Q_1} \cdot (Q_3 \cdot Q_2) \cdot clk^{\beta}, \tag{22}$$

$$Q_1^{\alpha} + Q_1^{\beta} = (Q_3 \cdot Q_2 \cdot \overline{Q_0} + Q_0) \cdot clk^{\beta} = Q_0^{\beta} + \overline{Q_0} \cdot (Q_3 \cdot Q_2) \cdot clk^{\beta}, \tag{23}$$

$$Q_0^{\alpha} + Q_0^{\beta} = clk^{\beta}. \tag{24}$$

Based on (2) and (4), (22) and (23) can be re-expressed as

$$Q_2^{\alpha} + Q_2^{\beta} = [Q_1 + (Q_3 \cdot Q_2) \cdot clk]^{\beta}, \tag{25}$$

$$Q_1^{\alpha} + Q_1^{\beta} = [Q_0 + (Q_3 \cdot Q_2) \cdot clk]^{\beta}, \quad (26)$$

Obviously, if we take $clk_3 = Q_2$, $clk_2 = [Q_1 + (Q_3 \cdot Q_2) \cdot clk]$, $clk_1 = [Q_0 + (Q_3 \cdot Q_2) \cdot clk]$ and $clk_0 = clk$, the covering relation will set the excitation functions of all the four flip-flops as $D_i = \overline{Q_i}$ (i =0,1,2,3). On the other hand, if we use the master clock for triggering all four flip-flops, we obtain the following complicated excitation functions:

$D_3 = Q_2 \cdot Q_1 \cdot Q_0 + Q_3 \cdot \overline{Q_2}$,

$D_2 = \overline{Q_2} \cdot Q_1 \cdot Q_0 + \overline{Q_3} \cdot \overline{Q_1} + \overline{Q_3} \cdot \overline{Q_0}$,

$D_1 = Q_3 \cdot Q_2 + \overline{Q_1} \cdot Q_0 + Q_1 \cdot \overline{Q_0}$,

$D_0 = \overline{Q_0}$.

Since the above D_3, D_2 and D_1 have complicated forms their corresponding synchronous circuit realization will have a complicated combinational circuit with more node capacitance and hence higher power dissipation. On the other hand, the corresponding asynchronous circuit realization with $D_i = \overline{Q_i}$ is much simpler. There is power saving since the four flip-flops are isolated from the triggering clock in the idle cycles.

IV. SYNCHRONOUS DERIVED CLOCK AND ITS APPLICATION

In the Example 1 of the last section we take $clk_i^{\beta} = Q_0^{\beta}$, ($i$ =1,2,3). From (12) we can also write clk_i^{β} as $clk_i^{\beta} = Q_0 \cdot clk^{\beta}$, ($i$ =1,2,3). Comparing this with (4), we have $g_i = 0$, $p_i = Q_0$ and $clk_i = Q_0 \cdot clk$. According to this form of the derived clock we get another asynchronous design, as shown in Fig.4(a). At the first glance, the circuit has one AND gate more than the design in Fig.3(b). Besides, it appears that the derived clock $clk_{1\text{-}3}$ may have an increased phase delay. However, the timing relation shown in Fig.1 indicates that the transition delay of $clk_{1\text{-}3}$ is independent of the delay of the Q_0 output. The delay between clk and $clk_{1\text{-}3}$ is only $2t_g$ (t_g is the average delay of a gate), which is less than the delay of the flip-flop output.

Figure 4. BCD code up-counter by gating clock
(a) asynchronous design, (b) synchronous design

Based on the above discussion, we can rewrite $clk_i = Q_0 \cdot clk$ as $clk_i^{*} = \overline{\overline{Q_0} + \overline{clk}}$. Besides, we take $\overline{clk}$ from the previous stage of the clock tree. Thus, we obtain a new design, as shown in Fig.4(b). If we consider delay of the

inverter and NOR gate being roughly the same, the falling transitions of *clk* and clk^{*}_{1-3} in the circuit will occur simultaneously. This design is synchronous in the sense that all flip-flops are triggered in synchrony with the global clock.

We simulated the new design in Fig.4(b) by SPICE 3f3 using 2μ CMOS technology, which proved that the new design has an ideal logic operation. We also measure the power dissipation of two synchronous designs in Fig.3(a) and Fig.4(b). The power dissipation diagrams are shown in Fig.5, and prove that the new design reduces the power dissipation by 22%.

Figure 5. Power dissipation diagram

V. CONCLUSION

The behavioral description of a clock is the basis to analyze its triggering action on flip-flops. Based on it, two types of clock-gating were introduced to form a derived clock. We showed that the procedure for designing a derived clock could be systematized so as to isolate the triggered flip-flop from the master clock in its idle cycles. The achieved power saving can be significant. However, the additional clock skew may lower the maximum operation frequency. Based on analyzing the timing relation in clock-gating, we then presented a new technique for generating the derived clock, which is synchronous with the master clock. Circuit simulation proved the quality of the new derived clock and its capability to reduce power dissipation. The engineering issues mentioned in (3) have thus been resolved for practical application, opening the path for widespread adoption of the clock-gating technique in low power design of custom ICs.

REFERENCES

1. M. Pedram, “Power minimization in IC Design: Principles and applications,” *ACM Transactions on Design Automaton*, vol. 1, no. 1, pp.3-56, Jan. 1996.

2. G. Friedman, “Clock distribution design in VLSI circuits: an overview,” in *Proc. IEEE ISCAS*, San Jose, pp.1475-1478, May 1994.

3. E. Tellez, A. Farrah and M. Sarrafzadeh, “Activity-driven clock design for low power circuits,” in *Proc. IEEE ICCAD*, San Jose, pp.62-65, Nov. 1995.

Fig.1 Timing relationship of *clk*, $p_i(g_i)$, p_i·clk and p_i+clk

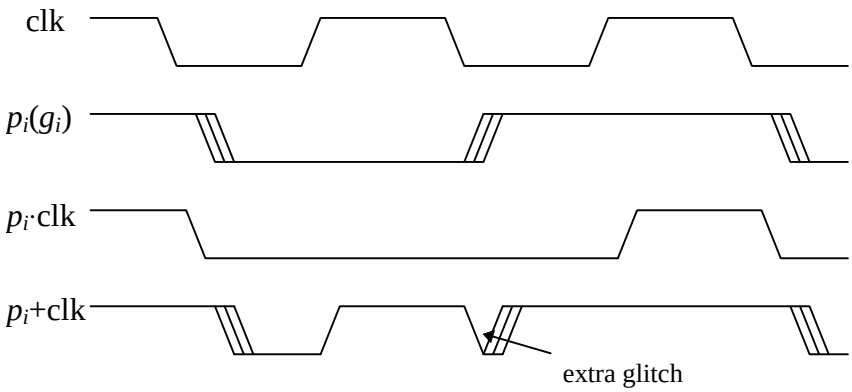
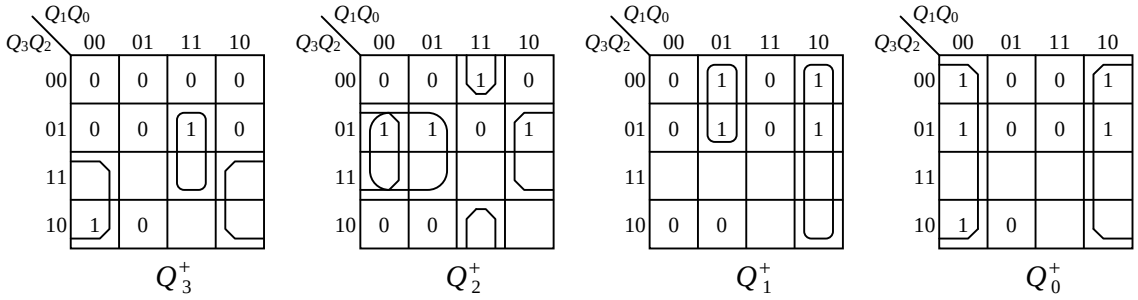
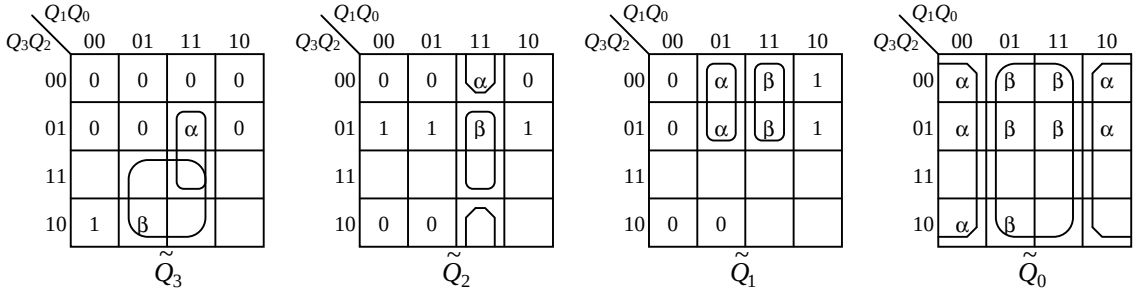


Fig.2. (a) Next state Karnaugh maps, (b) behavior Karnaugh maps, (c) simplified next state Karnaugh maps

(a)



(b)



(c)

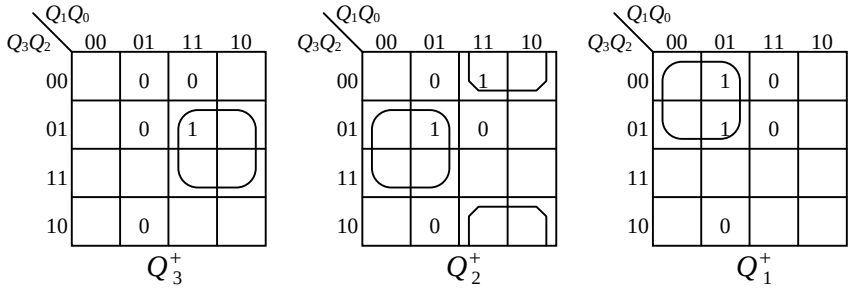


Fig 3. Circuit realizations of BCD code up-counter (a) synchronous design, (b) asynchronous design

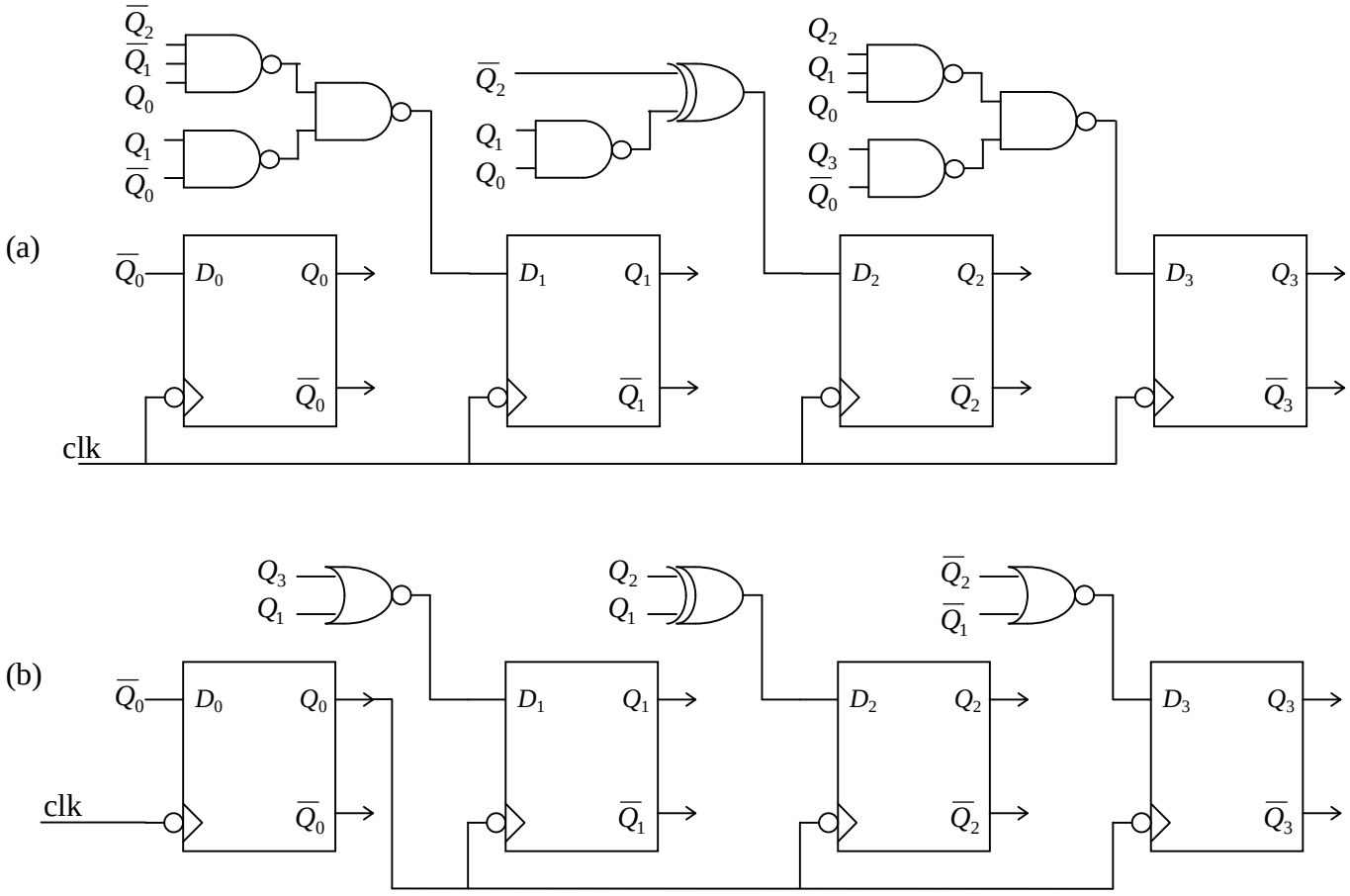


Fig 4. BCD code up-counter by gating clock (a) asynchronous design, (b) synchronous design

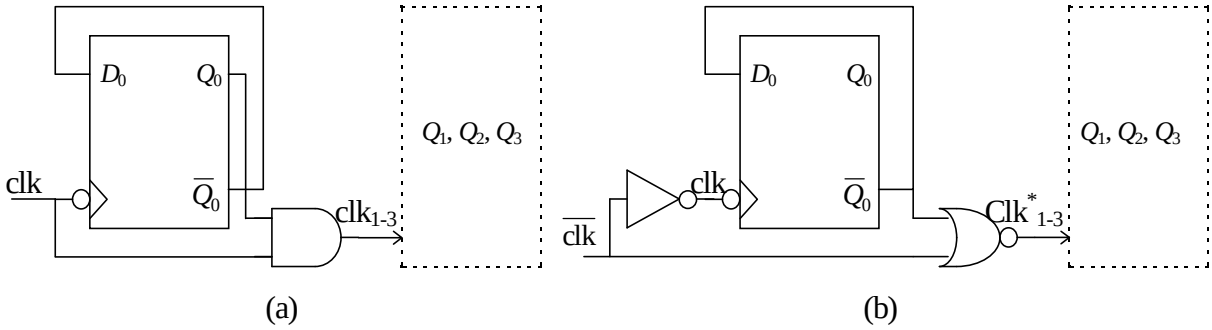


Fig. 5. Power dissipation diagram