Inverter Switching characteristics

7.6. With the parameters shown, a symmetrical design with $\beta_n = \beta_p$ gives $V_M = (V_{DD}/2) = 1.5$ V. Increasing the ratio to $(\beta_n/\beta_p) = 1.5$ gives $V_M = 1.42$ V, while $(\beta_n/\beta_p) = 2.5$ decreases the midpoint voltage to $V_M = 1.31$ V. It is also possible to use a ratio of $(\beta_n/\beta_p) < 1$, which shifts the VTC toward the right, i.e., $V_M > (V_{DD}/2)$. However, this is rarely used since the pFET aspect ratios get quite large.

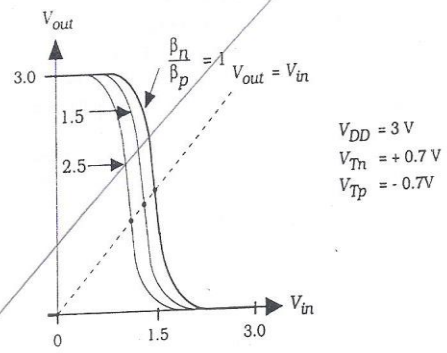


Figure 7.6 Dependence of V_M on the device ratio

7.2 Inverter Switching Characteristics

High-speed digital system design is based on the ability to perform calculations very quickly. This requires that logic gates introduce a minimum amount of time delay when the inputs change. Designing fast logic circuits is one of the more challenging (but critical) aspects of VLSI physical design. As with the DC analysis, analyzing the NOT gate provides a basis for studying more complicated circuits.

The general features of the problem are shown in Figure 7.7. An input voltage $V_{in}(t)$ is applied to the inverter, resulting in an output voltage $V_{out}(t)$. We assume that $V_{in}(t)$ has step-like characteristics and makes an abrupt transition from 0 to 1 (i.e., to a voltage of V_{DD}) at time t_1, and back down to 0 at time t_2. The output waveform reacts to the input, but the output voltage cannot change instantaneously. The output 1-to-0 transition introduces a **fall time** delay of t_f, while the 0-to-1 change at the output is described by the **rise time** t_r. The rise and fall times can be calculated by analyzing the electronic transitions of the circuits.

The rise and fall time delays are due to the parasitic resistance and

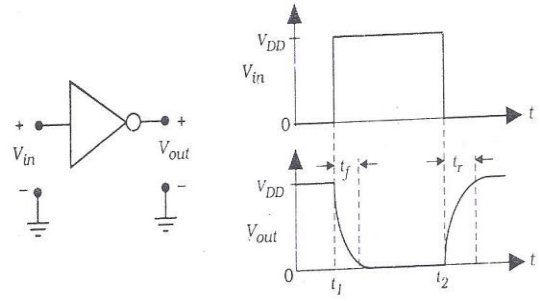


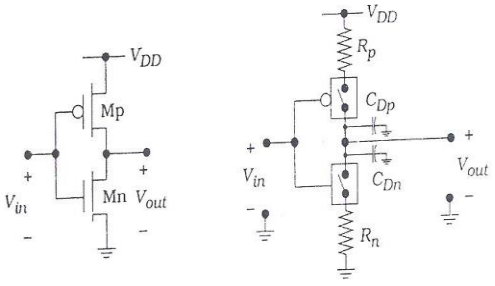
Figure 7.7 General switching waveforms

capacitances of the transistors. Consider the NOT circuit shown in Figure 7.8(a). Both FETs can be replaced by their switch equivalents, which results in the simplified RC model in Figure 7.8(b). It is worth recalling that the actual values of the components depend upon the device dimensions. Once we specify the aspect ratios $(W/L)_n$ and $(W/L)_p$, we can calculate R_n and R_p using

$$R_n = \frac{1}{\beta_n(V_{DD} - V_{Tn})} \qquad (7.28)$$

$$R_p = \frac{1}{\beta_p(V_{DD} - |V_{Tp}|)}$$

Knowing the layout dimensions of each FET allows us to find the capacitances C_{Dn} and C_{Dp} at the output node. The formulas are given by



(a) FET circuit (b) RC switch model equivalent

Figure 7.8 RC switch model equivalent for the CMOS inverter

$$C_{Dn} = C_{GSn} + C_{DBn} = \frac{1}{2}C_{ox}LW_n + C_{jn}A_n + C_{jswn}P_n \qquad (7.29)$$

$$C_{Dp} = C_{GSp} + C_{DBp} = \frac{1}{2}C_{ox}LW_p + C_{jp}A_p + C_{jswp}P_p$$

where we have added n and p subscripts to specify the nFET or pFET quantities, respectively.[1] It is significant to remember that increasing the channel width of a FET increases the parasitic capacitance values.

There is one more important point that needs to be included before we obtain a complete model. In a logic chain, every logic gate must drive another gate, or set of gates, to be useful. The number of gates is specified by the **fan-out** (FO) of the circuit. The fan-out gates act as a **load** to the driving circuit because of their **input capacitance** C_{in}. Consider the inverter shown in Figure 7.9(a). The input capacitance of the inverter is just the sum of the FET capacitances

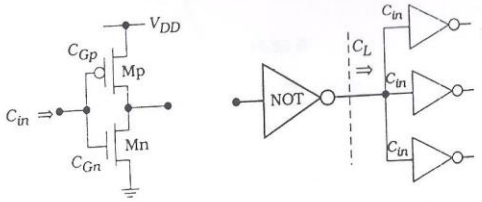
$$C_{in} = C_{Gp} + C_{Gn} \qquad (7.30)$$

Figure 7.8(b) shows the effect of input capacitance for a fan-out of FO = 3. The input capacitance to each gate acts as an **external load capacitance** C_L to the driving gate. In this example, it is easily seen that

$$C_L = 3C_{in} \qquad (7.31)$$

is the value of the load presented to the NOT gate.

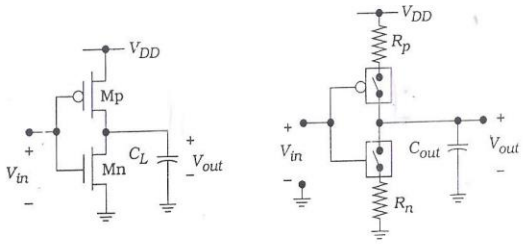
We may now calculate the switching times of the inverter. Figure 7.10 illustrates the general problem. A CMOS NOT gate is used to drive an external load capacitance C_L as in Figure 7.10(a). This gives the complete



(a) Single stage (b) Loading due to fan-out

Figure 7.9 Input capacitance and load effects

[1] Note that the source capacitances C_{Sp} and C_{Sn} do not enter the problem as they are at the power supply and ground, respectively, and have constant voltages.



(a) External load (b) Complete switching model

Figure 7.10 Evolution of the inverter switching model

switching model shown in Figure 7.10(b) where the total output capacitance is defined as

$$C_{out} = C_{FET} + C_L \qquad (7.32)$$

The FET capacitances shown earlier in Figure 7.8 have been merged into the single term

$$C_{FET} = C_{Dn} + C_{Dp} \qquad (7.33)$$

and are the parasitic internal contributions that cannot be eliminated. These add with C_L since all elements are in parallel. The total output capacitance C_{out} is the load that the gate must drive; the numerical value varies with the load.

Example 7.2

Let us apply this analysis to find the capacitances in the NOT gate shown in Figure 7.11. It is assumed that all dimensions have units of microns (μm).

First we will find the gate capacitances using

$$C_{Gp} = (2.70)(1)(8) = 21.6 \text{ fF}$$
$$C_{Gn} = (2.70)(1)(4) = 10.8 \text{ fF} \qquad (7.34)$$

Next, note that the overlap distance L_o is specified as 0.1 μm, which should be included in the area and perimeter factors in the junction capacitances. For the pFET, the p+ capacitance is

$$C_p = C_j A_{bot} + C_{jsw} P_{sw} \qquad (7.35)$$

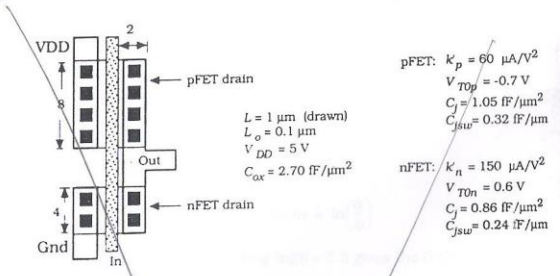


Figure 7.11 Example of capacitance calculations

so

$$C_p = (1.05)(8)(2.1) + (0.32)2(8 + 2.1) = 24.10 \text{ fF} \tag{7.36}$$

The total capacitance at the pFET drain is therefore given by

$$C_{Dp} = \frac{21.6}{2} + 24.10 = 34.9 \text{ fF} \tag{7.37}$$

The nFET drain is analyzed using the same approach. The n+ junction capacitance is

$$C_n = (0.86)(4)(2.1) + (0.24)(2)(4 + 2.1) = 10.15 \text{ fF} \tag{7.38}$$

so that

$$C_{Dn} = \frac{10.8}{2} + 10.15 = 15.55 \text{ fF} \tag{7.39}$$

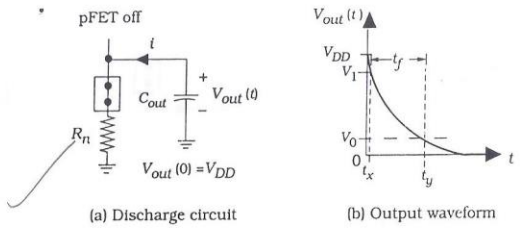
is the total capacitance at the drain of the nFET. Adding gives

$$\begin{aligned} C_{FET} &= C_{Dp} + C_{Dn} \\ &= 34.9 + 15.55 \\ &= 50.45 \text{ fF} \end{aligned} \tag{7.40}$$

as the total internal FET capacitance. The total capacitance at the output is

$$C_{out} = 50.45 + C_L \tag{7.41}$$

in fF, where C_L is the external load (also in fF).



(a) Discharge circuit (b) Output waveform

Figure 7.12 Discharge circuit for the fall time calculation

7.2.1 Fall Time Calculation

Let us start by calculating the output fall time t_f. We will shift the time origin such that V_{in} changes from 0 to V_{DD} at time $t = 0$. The initial condition at the output is $V_{out}(0) = V_{DD}$. When the input is switched, the nFET goes active while the pFET is driven into cutoff. In terms of the switch models, the nFET switch is closed and the pFET switch is open. This gives us the simplified discharge circuit shown in Figure 7.12(a). The capacitor C_{out} is initially charged to a voltage V_{DD}, and is allowed to discharge to 0 V through the nFET resistance R_n. The current leaving the capacitor is

$$i = -C_{out}\frac{dV_{out}}{dt} = \frac{V_{out}}{R_n} \tag{7.42}$$

which gives the differential equation for the discharge event. Solving with the initial condition $V_{out}(0) = V_{DD}$ results in the well-known form

$$V_{out}(t) = V_{DD}e^{-t/\tau_n} \tag{7.43}$$

where

$$\tau_n = R_n C_{out} \tag{7.44}$$

is the nFET **time constant** with units of seconds. The function is plotted in Figure 7.12(b).

The fall time is traditionally defined to be the time interval from $V_1 = 0.9\,V_{DD}$ to $V_0 = 0.1\,V_{DD}$, which are respectively known as the 90% and the 10% voltages as referenced to the full rail swing of V_{DD}. Rearranging the solution to the form

$$t = \tau_n \ln\left(\frac{V_{DD}}{V_{out}}\right) \tag{7.45}$$

allows us to calculate the time t needed to fall to a particular voltage V_{out}. From the drawing we see that

$$\begin{aligned} t_f &= t_y - t_x \\ &= \tau_n \ln\left(\frac{V_{DD}}{0.1V_{DD}}\right) - \tau_n \ln\left(\frac{V_{DD}}{0.9V_{DD}}\right) \\ &= \tau_n \ln(9) \end{aligned} \tag{7.46}$$

where we have used the identity

$$\ln(a) - \ln(b) = \ln\left(\frac{a}{b}\right) \tag{7.47}$$

in the last step. Approximating $\ln(9) \approx 2.2$ gives the final result

$$t_f \approx 2.2\tau_n \tag{7.48}$$

as the fall time for the circuit. The output fall time in a generic digital logic gate is usually called the output **high-to-low time** t_{HL} and is identical to the value computed here:

$$t_{HL} = t_f \tag{7.49}$$

The two symbols will be used interchangeably in the discussion.

7.2.2 The Rise Time

The rise time calculation follows in the same manner. Initially, the input voltage is at $V_{in} = V_{DD}$ and is switched to $V_{in} = 0$ V; we time shift this event to occur at $t = 0$ for simplicity. This turns on the pFET while simultaneously driving the nFET into cutoff, so that the simplified charging circuit of Figure 7.13(a) is valid. The output voltage at $t = 0$ is given by $V_{out}(0) = 0$ V.

The charging current is given by

$$i = C_{out}\frac{dV_{out}}{dt} = \frac{V_{DD} - V_{out}}{R_p} \tag{7.50}$$

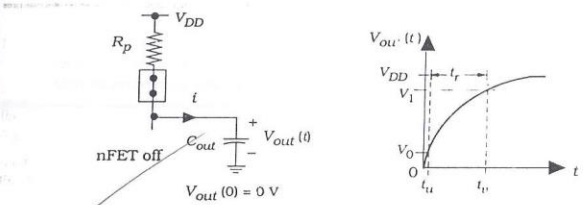
Solving and applying the initial condition gives the exponential form

$$V_{out}(t) = V_{DD}[1 - e^{-t/\tau_p}] \tag{7.51}$$

where the pFET time constant is defined by

$$\tau_p = R_p C_{out} \tag{7.52}$$

Figure 7.13(b) shows the output voltage as a function of time. The rise time is taken between 10% and 90% points such that



(a) Charge circuit (b) Output waveform

Figure 7.13 Rise time calculation

$$t_r = t_v - t_u \tag{7.53}$$

A little algebra yieids the expression

$$t_r = \ln(9)\tau_p \approx 2.2\tau_p \tag{7.54}$$

for the rise time t_r. This has the same form as the fall time t_f because of the symmetry of the charge and discharge circuits. The rise time is identical to the output **low-to-high time** t_{LH}; the symbols will be used interchangeably.

The low-to-high time t_{LH} and the high-to-low time t_{HL} represent the shortest amount of time needed for the output to change from a logic 0 to logic 1 voltage, or from a logic 1 to a logic 0 voltage, respectively. Let us assume that the input is a square wave with a period of T sec such that the voltage is 0 for $(T/2)$ and V_{DD} for a $(T/2)$ time interval.[2] We then define the **maximum signal frequency** as

$$f_{max} = \frac{1}{t_{HL} + t_{LH}} = \frac{1}{t_r + t_f} \tag{7.55}$$

since this is the largest frequency that can be applied to the gate and still allow the output to settle to a definable state.[3] If the signal frequency exceeds f_{max} the output voltage of the gate will not have sufficient time to stabilize to the correct value.

---

[2] This defines what is known as a 50% duty cycle.

[3] This definition assumes that t_{HL} and t_{LH} have the same order of magnitude to be useful.

Example 7.3

Consider an inverter circuit that has FET aspect ratios of $(W/L)_n = 6$ and $(W/L)_p = 8$ in a process where

$$k'_n = 150\ \mu A/V^2 \qquad V_{Tn} = +0.70\ V$$
$$k'_p = 62\ \mu A/V^2 \qquad V_{Tp} = -0.85\ V \tag{7.56}$$

and uses a power supply voltage of $V_{DD} = 3.3$ V. The total output capacitance is estimated to be $C_{out} = 150$ fF. Let us compute the rise and fall times using the equations derived above.

Consider first the fall time. The pFET resistance is given by

$$R_p = \frac{1}{\beta_p(V_{DD} - |V_{Tp}|)} = \frac{1}{(62\times10^{-6})(8)(3.3-0.85)} = 822.9\ \Omega \tag{7.57}$$

The time constant for the charging event is computed using the RC product R_pC_{out} to find

$$\tau_p = (822.9)(150\times10^{-15}) = 123.43 \text{ ps} \tag{7.58}$$

where 1 ps (picosecond) is 10^{-12} sec. The rise time is

$$t_r = 2.2\tau_p = 271.55 \text{ ps} \tag{7.59}$$

The fall time is calculated in a similar manner. First, we find the nFET resistance

$$R_n = \frac{1}{\beta_n(V_{DD} - V_{Tn})} = \frac{1}{(150\times10^{-6})(6)(3.3-0.70)} = 427.35\ \Omega \tag{7.60}$$

so that the discharge time constant is

$$\tau_p = (427.35)(150\times10^{-15}) = 64.1 \text{ ps} \tag{7.61}$$

The fall time is

$$t_f = 2.2\tau_n = 141.0 \text{ ps} \tag{7.62}$$

Combining these results, the maximum signal frequency is

$$f_{max} = \frac{1}{t_r + t_f} = \frac{1}{(271.55 + 141.0)\times10^{-12}} = 2.42 \text{ GHz} \tag{7.63}$$

where 1 GHz = 10^9 Hz. Although this is a very high frequency, it is important to remember that this refers only to a single inverter.

7.2.3 The Propagation Delay

The propagation delay time t_p is often used to estimate the "reaction" delay time from input to output. When we use step-like input voltages, the propagation delay is defined by the simple average of the two time intervals shown in Figure 7.14 by

$$t_p = \frac{(t_{pf} + t_{pr})}{2} \tag{7.64}$$

In this expression, t_{pf} is the output fall time from the maximum level to the "50%" voltage line, i.e., from V_{DD} to $(V_{DD}/2)$; t_{pr} is the propagation rise time from 0 V to $(V_{DD}/2)$. Using the exponential equations for V_{out} we obtain

$$t_{pf} = \ln(2)\tau_n \qquad t_{pr} = \ln(2)\tau_p \tag{7.65}$$

Approximating $\ln(2) \approx 0.693$ then gives

$$t_p \approx 0.35(\tau_n + \tau_p) \tag{7.66}$$

The propagation delay time is a useful estimate of the basic delay, but does not provide detailed information on the rise and fall times as individual quantities. Propagation delays are commonly used in basic logic simulation programs.

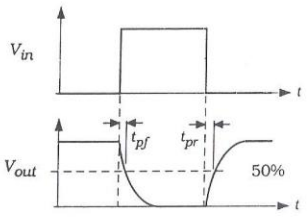


Figure 7.14 Propagation time definitions