

## CMOS Gate Delays

Assume we have a symmetrical CMOS inverter with

$$V_{Tn} = |V_{Tp}| \equiv V_T \quad \text{and} \quad K_n = K_p.$$

Also assume that both the n- and p-channel devices are minimum gate length devices, so

$$L_n = L_p = L_{\min},$$

and that

$$C_{o\ xn}^* = C_{o\ xp}^* \equiv C_{o\ x}^*.$$

Finally, assume  $\mu_e = 2 \mu_h$ , so if we make the p-channel device twice as wide as the n-channel device, we get the desired K equality; i.e.,

$$W_p = 2 W_n,$$

yields

$$K_p = \frac{W_p}{L_p} \mu_h C_{o\ xp}^* = \frac{2 W_n}{L_n} \frac{\mu_e}{2} C_{o\ xn}^* = K_n.$$

The gate delay of an inverter is the sum of the times it takes the gate to switch from a LO to a HI output, and from a HI to a LO output. To estimate these times for a CMOS gate we first note that during the LO to HI cycle, the load capacitance,  $C_L$ , is charged from 0 V to  $V_{DD}$ , which requires a total charge of  $C_L V_{DD}$ , through the p-channel device. During much of this cycle the p-channel MOSFET will be in saturation with  $V_{GS} = -V_{DD}$ , and so the current through it into the load will be

$$I_L = \frac{K_p}{2} (-V_{DD} + V_T)^2 = \frac{K_p}{2} (V_{DD} - V_T)^2$$

As the load charges up, and the p-channel device comes out of saturation, this current will decrease, but for sake of estimation, assume that it stays constant over the entire charging cycle. The charging time will then be the total charge divided by this current:

$$t_{LO-HI} = \frac{Q}{I} = \frac{2 C_L V_{DD}}{K_p (V_{DD} - V_T)^2}$$

During the HI to LO swing, this charge is discharged to ground through the n-channel MOSFET. Again the MOSFET will be in saturation initially with  $V_{GS} = V_{DD}$ , and

$$I_D = \frac{K_n}{2} (V_{DD} - V_T)^2$$

The current will decrease once  $v_{OUT}$  drops below  $(V_{DD} - V_T)$  and the transistor is no longer in saturation, but we can again estimate  $t_{HI-LO}$  by assuming it stays constant. We find

$$t_{HI-LO} = \frac{2 C_L V_{DD}}{K_n (V_{DD} - V_T)^2}$$

which is exactly the same since  $K_n = K_p$ . This symmetry is important because it gives us the shortest overall cycle time, and is another reason to make the K's equal. The total gate delay is the sum of these two times

$$t_{GD} = \frac{4 C_L V_{DD}}{K_n (V_{DD} - V_T)^2} = \frac{12 n}{\mu_e} L_{min}^2 \frac{V_{DD}}{(V_{DD} - V_T)^2}$$

To obtain the right-hand expression we have replaced  $K_n$  with  $(W_n/L_n) \mu_e C_{ox}^*$ , and  $C_L$  with  $n(W_p L_p + W_n L_n) C_{ox}^*$ , which is in turn  $3n W_n L_n C_{ox}^*$ . Written this way we can readily see the advantage of making  $L_{min}$  smaller.