I. CMOS Inverter: Propagation Delay

A. Introduction

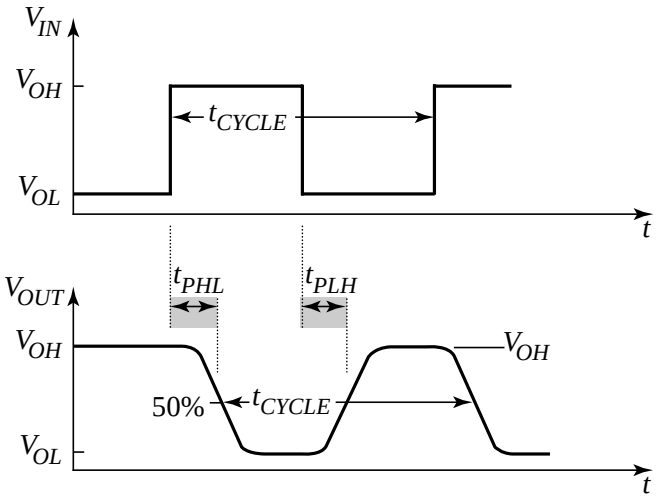
- Propagation delays t_{PHL} and t_{PLH} define ultimate speed of logic
- Define Average Propagation Delay

$$t_p = \frac{\left(t_{PHL} + t_{PLH}\right)}{2}$$

- Typical complex system has 20-50 propagation delays per clock cycle.
- Typical propagation delays < 1nsec

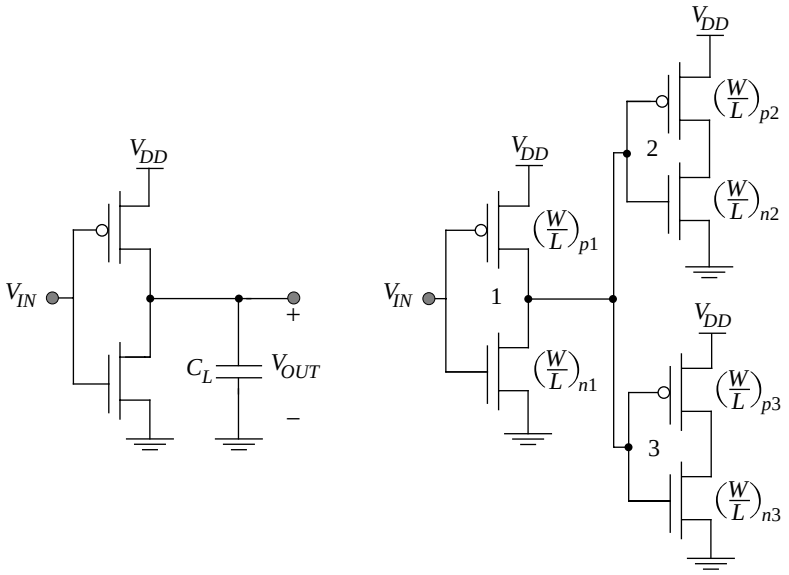
B. Hand Calculation

- Use an input signal that has t_r=0 and t_f=0 for hand calculation
- Calculate current drive
- Calculate capacitance being driven



C. Load Capacitance: Input to Next Stage

- Typical configuration: load capacitance C_L consists of the input capacitances of the next stage of inverters plus parasitic drain/bulk capacitance and **wiring** capacitance



(a) (b)

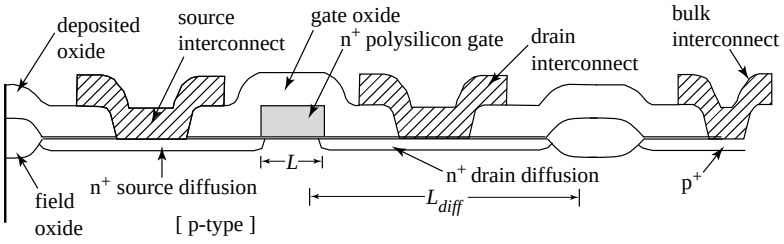
- Estimation of the input capacitance: n- and p-channel transistors in the next stage switch from triode through saturation to cutoff during a high-low or low-high transition
- Requires nonlinear charge storage elements to accurately model
- Hand Calculation use a rough estimate for an inverter

$$C_{in} = C_{ox}(W \cdot L)_p + C_{ox}(W \cdot L)_n$$

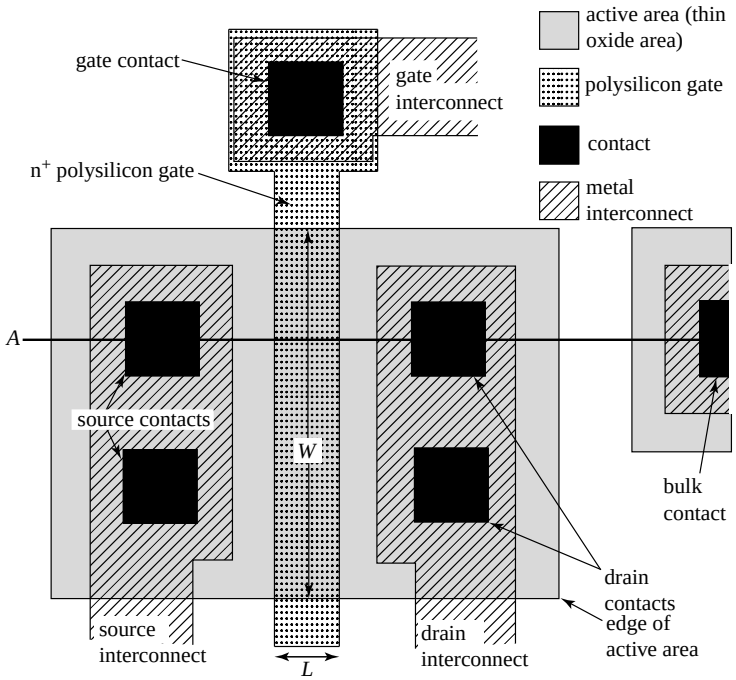
- C_G above is

$$C_G = C_{ox}(WL)_{p2} + C_{ox}(WL)_{n2} + C_{ox}(WL)_{p3} + C_{ox}(WL)_{n3}$$

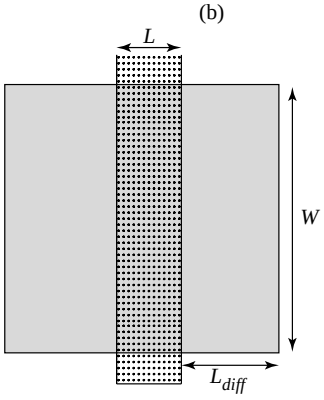
D. Parasitic Capacitance-Drain/Bulk Depletion



(a)



(b)



(c)

Calculation of Parasitic Depletion Capacitance

- Depletion $q_J(v_D)$ is non-linear --> take the worst case and use the zero-bias capacitance C_{jo} as a linear charge-storage element during the transient.
- "Bottom" of depletion regions of the inverter's drain diffusions contribute a depletion capacitance

$$C_{BOTT} = C_{Jn}(W_n L_{diffn}) + C_{Jp}(W_p L_{diffp})$$

- C_{Jn} and C_{Jp} being the zero-bias junction capacitances (fF/μm^2) for the n-channel MOSFET drain-bulk junction and the p-channel MOSFET drain-bulk junction, respectively.

Typical numbers: C_{Jn} and C_{Jp} are about 0.2 fF/μm^2

- "Sidewall" of depletion regions of the inverter's drain diffusions make an additional contribution:

$$C_{SW} = (W_n + 2L_{diffn})C_{JSWn} + (W_p + 2L_{diffp})C_{JSWp}$$

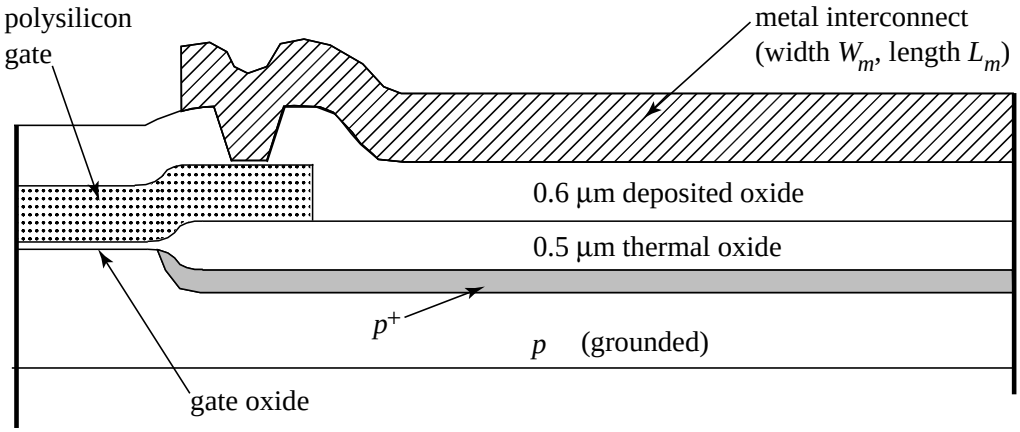
- C_{JSWn} and C_{JSWp} being the zero-bias sidewall capacitances (F/μm) for the n-channel MOSFET drain-bulk junction and the p-channel MOSFET drain-bulk junction, respectively.

Typical numbers: C_{JSWn} and C_{JSWp} are about 0.5 fF/μm

- The sum of C_{BOTT} and C_{SW} is the total depletion capacitance, C_{DB}

E. Parasitic Capacitance-Wires

- "Wires" consist of metal lines connecting the output of the inverter to the input of the next stage



- The p^+ layer (i.e., heavily doped with acceptors) under the thick thermal oxide (500 nm = 0.5 μm) and deposited oxide (600 nm = 0.6 μm) depletes only slightly when positive voltages appear on the metal line, so the capacitance is approximately the oxide capacitance:

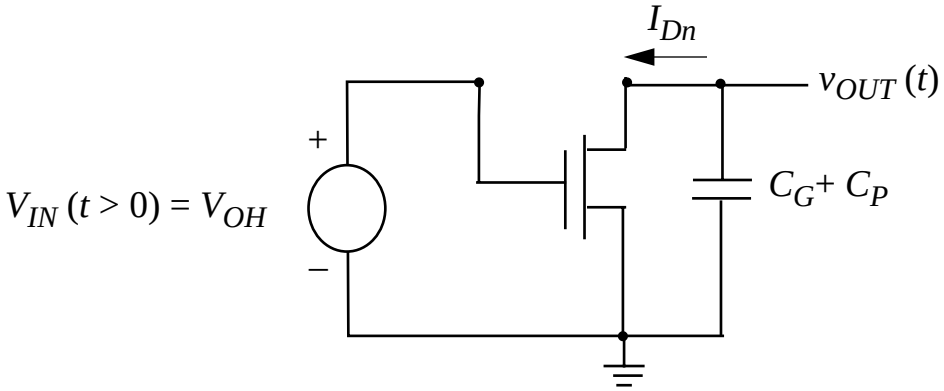
$$C_{wire} = C_{thickox}(W_m \cdot L_m)$$

- where the oxide thickness = 500 nm + 600 nm = 1.1 μm.
- For large digital systems, the parasitic capacitance can dominate the load capacitance

$$C_L = C_G + C_P = C_G + (C_{DB} + C_{wire})$$

F. High-to-Low Transition (at Output)

- Assume V_{IN} switches instantly from low to high.
- Driver transistor (n-channel) switches from cutoff to saturation
- p-channel load switches from triode to cutoff
- Circuit during high-to-low transition:



- The voltage on the load capacitor at $t = 0^-$ was V^+
- Since n-channel MOSFET is initially saturated and the input voltage is a constant, the drain current is initially $I_{Dn(sat)}$ for $V_{GS} = V^+$.

$$I_{Dn(sat)} = (1/2)\,\mu_n C_{ox} (W/L)_n (V_{OH} - V_{Tn})^2$$

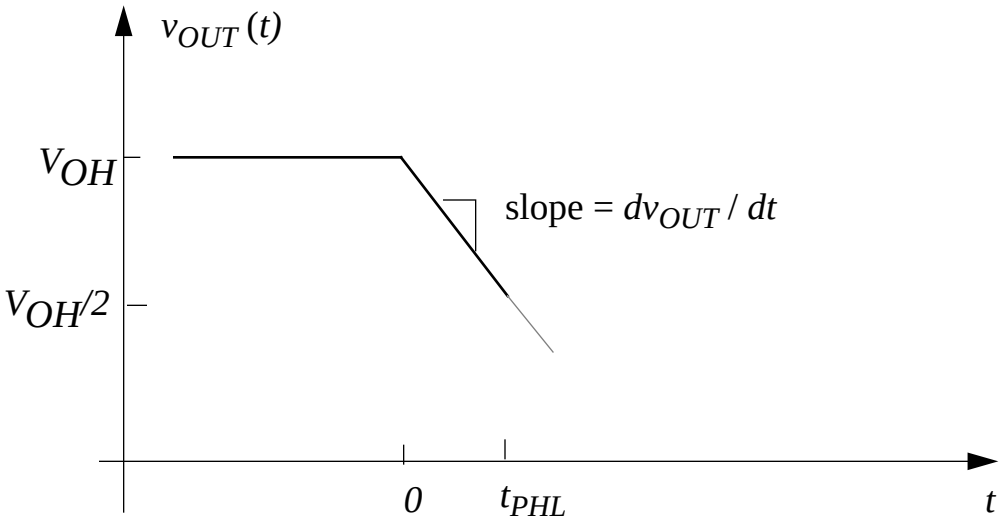
$$V_{OH} = V^+$$

- Propogation delay is (by definition) the time required for V_{OUT} to reach $V^+ / 2$:

$$\frac{dv_{OUT}}{dt} = \frac{d}{dt}\left(\frac{Q_L}{C_G + C_P}\right) = \frac{-I_{Dn(sat)}}{C_G + C_P}$$

G. Hand Calculation of t_{PHL}

- The output waveform is:



- The high-to-low propagation delay is given by

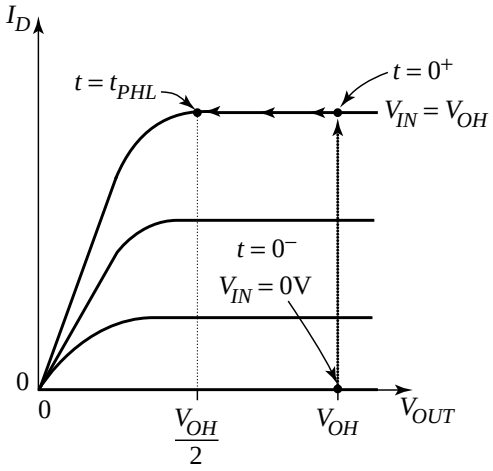
$$\frac{dv_{OUT}}{dt} = \frac{(V_{OH}/2) - V_{OH}}{t_{PHL}} = \frac{-I_{Dn(sat)}}{C_G + C_P}$$

- Solving for the delay:

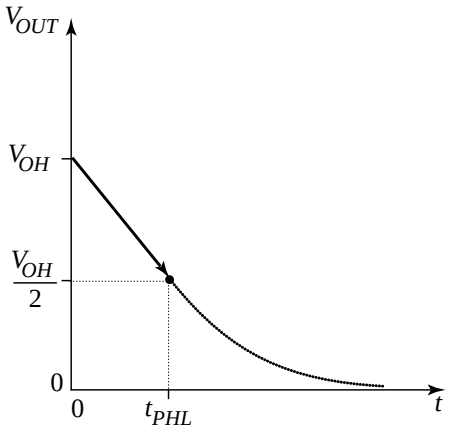
$$t_{PHL} = \frac{(C_G + C_P)\,(V_{OH}/2)}{(1/2)\,\mu_n C_{ox}\,(W/L)_n\,(V_{OH} - V_{Tn})^2}$$

H. Graphical Interpretation

- The n-channel driver remains saturated throughout the first half of the transition from high-to-low... all that matters, according to the definition of propagation delay for hand analysis.



(a)



(b)

I. Hand Calculation of t_{PLH}

- low-to-high transition, the p-channel load is supplying a constant current $-I_{Dp(sat)}$ to charge up the load and parasitic capacitance.

$$t_{PLH} = \frac{(C_G + C_P)\,(V_{OH}/2)}{(1/2)\,\mu_p C_{ox}\,(W/L)_p\,(V_{OH} + V_{Tp})^2}$$

- For identical propagation delays, the (W/L) of the p-channel load is a factor of two higher than for the n-channel driver, to compensate for the lower hole mobility in the channel.

II. Power Dissipation

- Energy from power supply needed to charge up the capacitor:

$$E_{charge} = \int V^{+} i(t)dt = V^{+}Q = (V^{+})^{2}(C_G + C_P)$$

- Energy stored in the capacitor:

$$E_{store} = \left(\frac{1}{2}\right)(C_G + C_P)\ (V^{+})^{2}$$

- Energy lost in p-channel MOSFET during charging:

$$E_{diss} = E_{charge} - E_{store} = \left(\frac{1}{2}\right)(C_G + C_P)\ (V^{+})^{2}$$

- During discharge, the n-channel MOSFET driver dissipates an identical amount of energy.
- *If* the charge/discharge cycle is repeated *f* times/second, where *f* is the clock frequency, the *dynamic power dissipation* is:

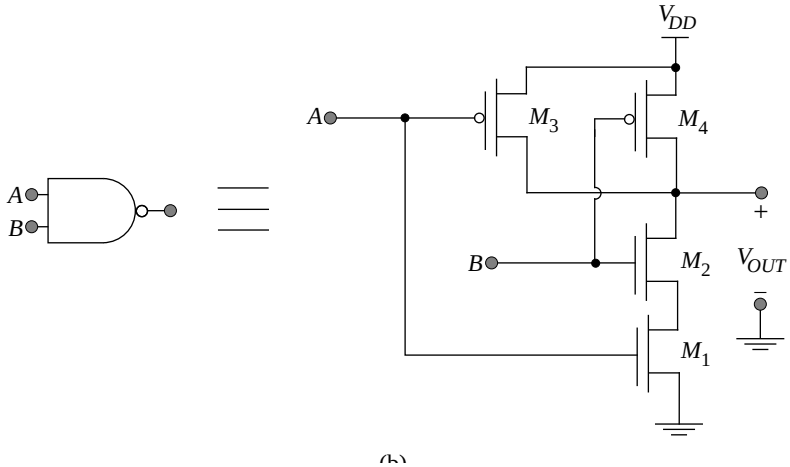
$$P = (2E_{diss}) \cdot f = (C_G + C_P)\ (V^{+})^{2} f$$

- In practice, many gates don't change state for every clock cycle, which lowers the power dissipation

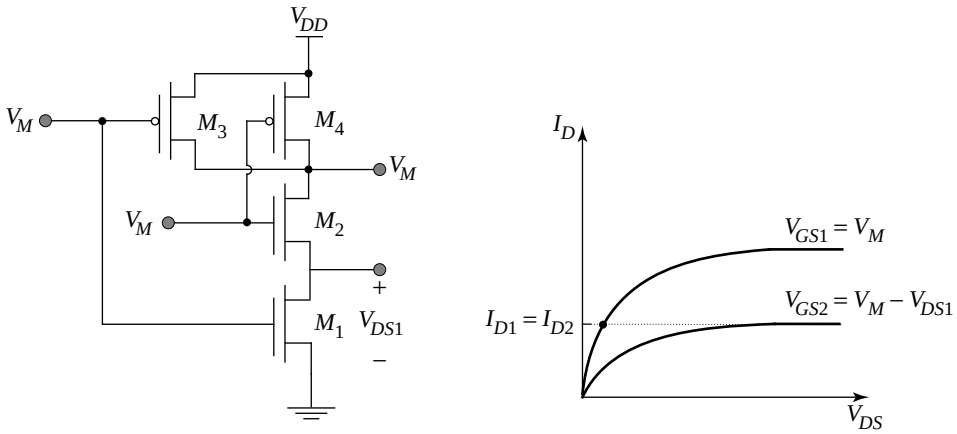
- Additional source of dissipation: power flow from V^{+} to ground when both transistors are saturated. Can be signifcant, but hard to estimate by hand. **Eliminate with Circuit Techniques.**

III. CMOS Static Logic Gates

A. CMOS Static NAND Gate

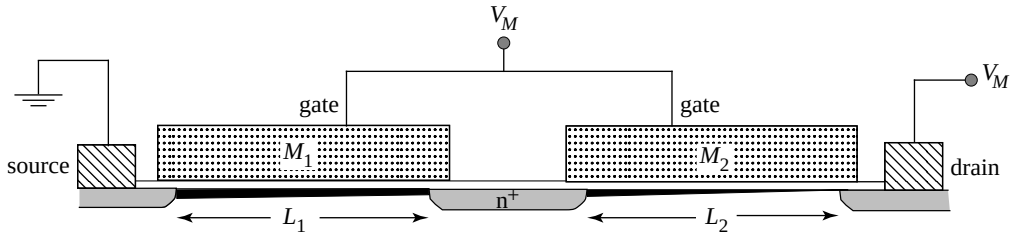


B. I-V Characteristics of n-channel devices

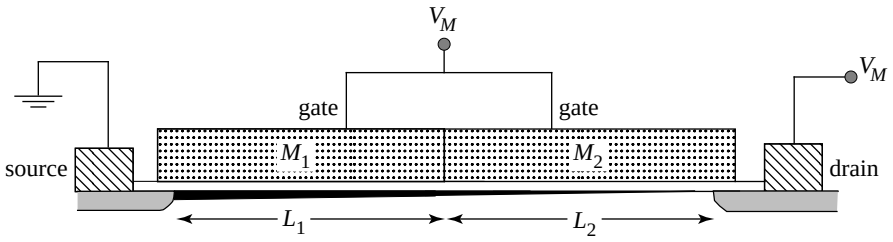


C. Define k_{neff} and k_{peff} for the NAND Gate

- Effective width of two p-channel devices is $2W_p$
- *BUT* worst case only 1 device is on...$k_{peff} = k_{p3} = k_{p4}$
- Effective length of two n-channel devices is $2L_n$ (Derivation in Text)
- $k_{neff} = k_{n1}/2 = k_{n2}/2$



(a)



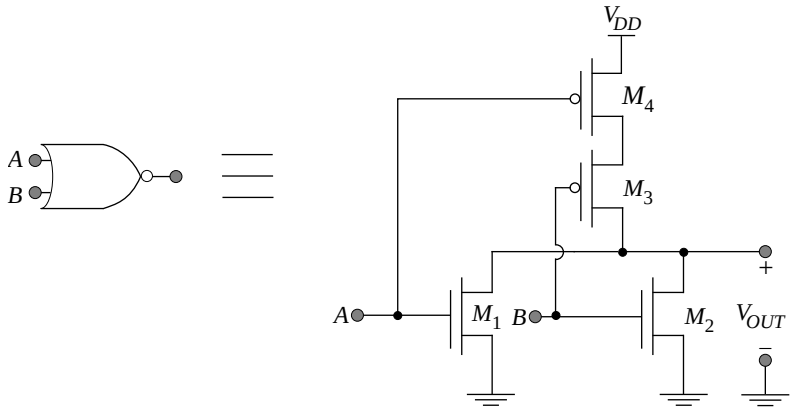
(b)

D. Calculation of static and transient performance

- $k_{peff} = k_{neff}$ is desirable (equal propagation delay; symmetrical transfer curve),
- Recall $\mu_n = 2\mu_p$ $\left(\frac{W}{L}\right)_n = \left(\frac{W}{L}\right)_p$ 2-input NAND Gate
- For an M-input NAND Gate

$$\left(\frac{W}{L}\right)_n = \frac{M}{2}\left(\frac{W}{L}\right)_p$$

E. CMOS Static NOR Gate



(a)

- Effective width of two n-channel devices is $2W_n$ ...$k_{neff} = 2k_{n1} = 2k_{n2}$
- *BUT* worst case only 1 device is on...$k_{neff} = k_{n1} = k_{n2}$
- Effective length of two p-channel devices is $2L_p$ (Same reasoning)
- $k_{peff} = k_{p3}/2 = k_{p4}/2$
- An *M*-input NOR gate-requires very wide p-channel devices since $\mu_p = \mu_n/2$

$$\left(\frac{W}{L}\right)_p = 2M\left(\frac{W}{L}\right)_n$$