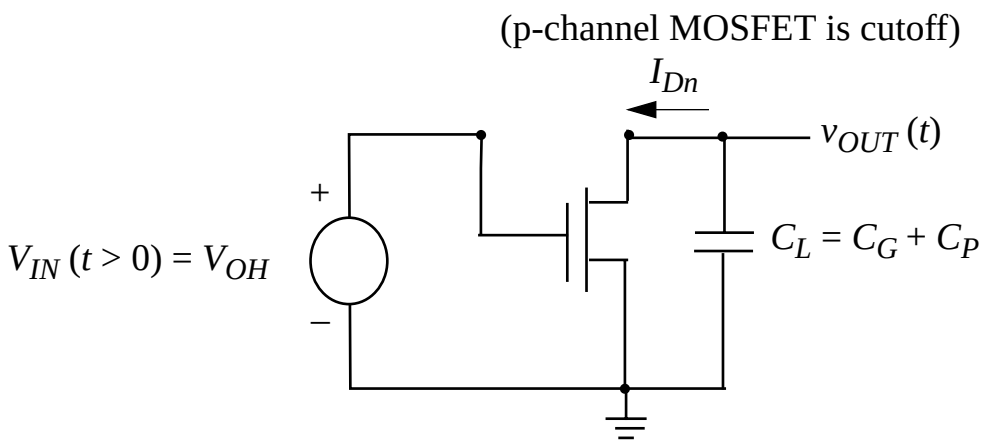
High-to-Low Propagation Delay t_{PHL}

- V_{IN} switches instantly from low to high. Driver transistor (n-channel) immediately switches from cutoff to saturation; the p-channel pull-up switches from triode to cutoff.
- Circuit during high-to-low transition:

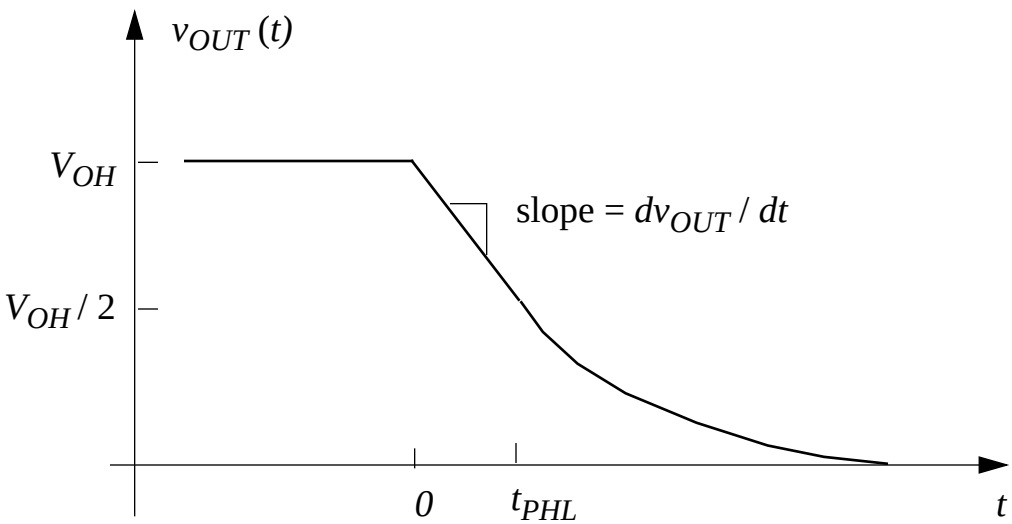


- The voltage on the load capacitor at $t = 0^-$ was V_{OH}; since n-channel MOSFET is saturated initially and the input voltage is a constant, the drain current is initially $I_{Dn(sat)}$ for $V_{GS} = V^+$.
- The high-to-low propagation delay t_{PHL} is (by definition) the time required for V_{OUT} to reach V_{OH} / 2:

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

Hand Calculation of t_{PHL}

- The output voltage decreases linearly over $0 < t < t_{PHL}$ if we assume that the MOSFET remains saturated:



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

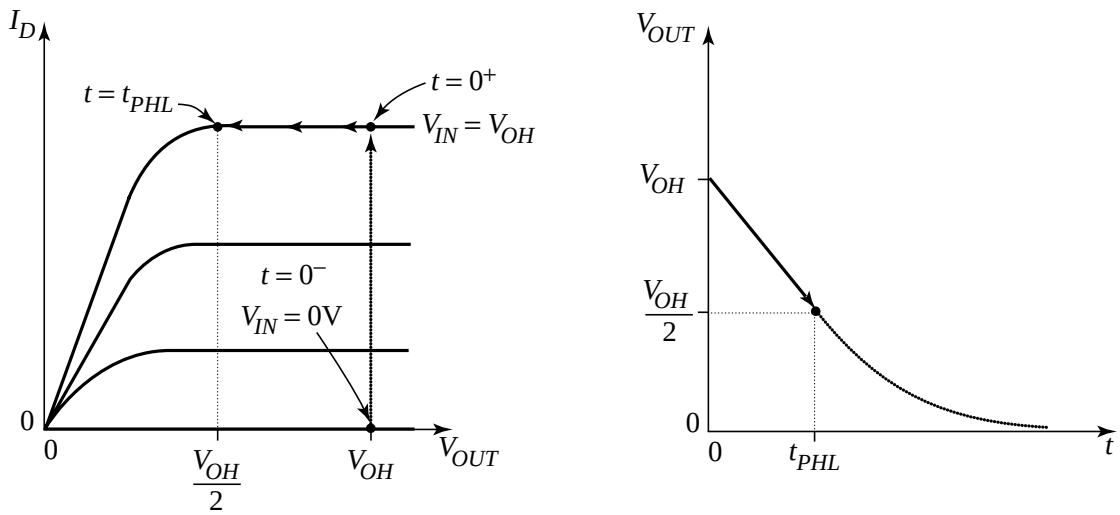
$$\frac{dv_{OUT}}{dt} = \frac{(V_{OH}/2) - V_{OH}}{t_{PHL}} = \frac{-I_{Dn(sat)}}{C_G + C_P}$$

Solving for the delay:

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

Graphical Interpretation

- The n-channel driver remains saturated throughout the first half of the transition from high-to-low...

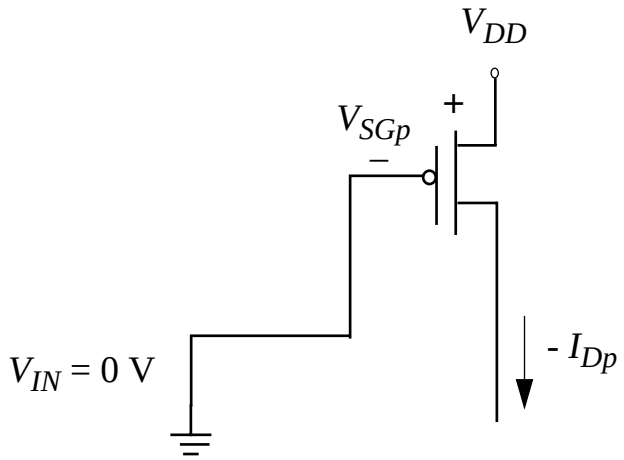


(a) (b)

note that the characteristics above are *not* for a square-law MOSFET, which would enter the triode region for $V_{OUT} < V_{OH} - V_{Tn}$; the error is not large enough to matter for hand calculations in any case

Low-to-High Propagation Delay t_{PLH}

- For the low-to-high transition, the n-channel device is cutoff and the p-channel MOSFET is initially saturated and supplying - $I_{Dp(sat)}$ to charge up the gate and parasitic capacitances.



- Therefore,

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

In order to have identical propagation delays, the width-to-length ratio of the p-channel pull-up must be twice that of the n-channel driver, in order to compensate for the lower hole mobility in the channel.

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:

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

In practice, many gates don't change state 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 significant, but hard to estimate by hand. Typical number: 25% of dynamic power dissipation.

Power (cont.)

- Practical numbers: C_L = 50 fF, f = 200 MHz, V^+ = 3 V, N_{gates} = 5 x 10^5

  P = 45 W ! (note that the real average depends on the average number switching per clock cycle)

Comparing Technologies -- the power-delay product

* Logic families are often compared by considering the product of the dynamic power dissipation and the propagation delay:

$$PDP = Pt_P \cong \left\{ (C_L + C_p)(V^+)^2 f \right\} \left\{ \frac{(C_L + C_p)(V^+/2)}{\frac{1}{2}k_N(V^+ - V_{Tn})^2} \right\}$$

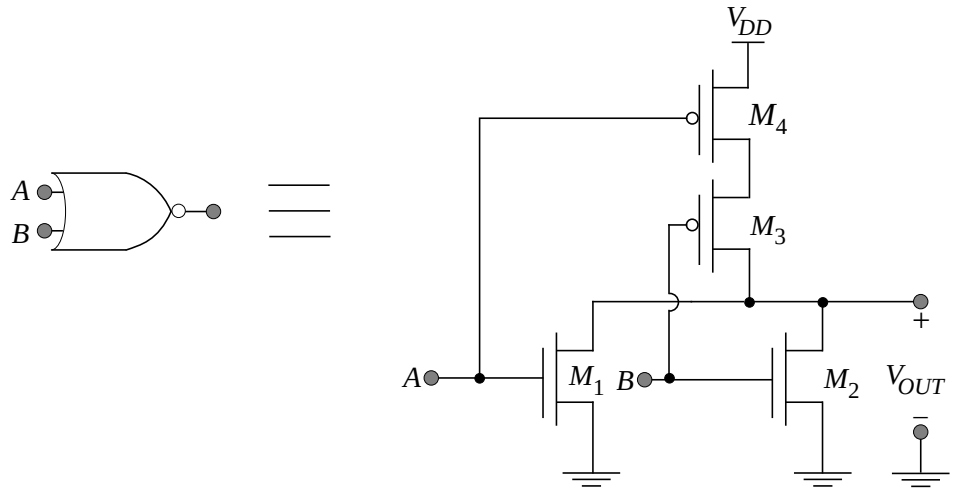
where V^+ has been substituted for V_{OH} to achieve a more universal result.

* For $V^+ >> V_{Tn}$,

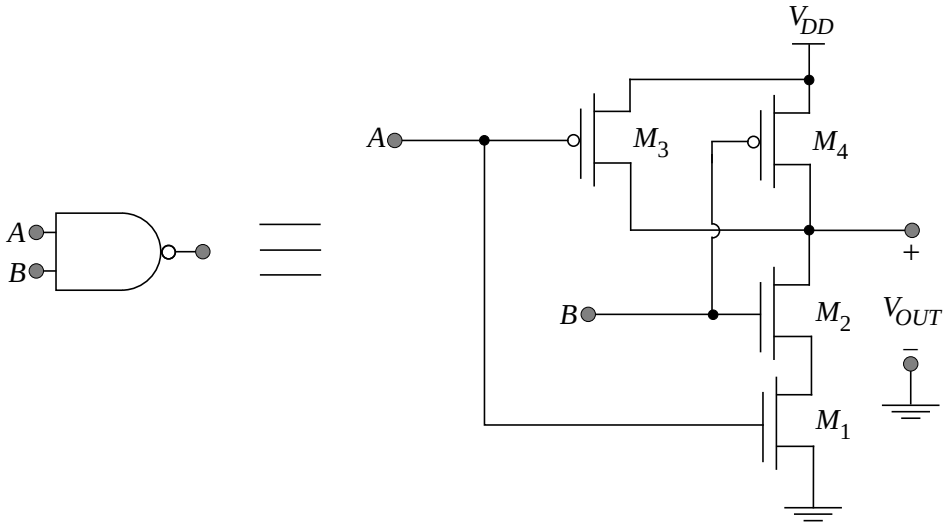
$$PDP \cong \frac{(C_L + C_p)^2 V^+ f}{k_N}$$

CMOS Static Logic Gates

- "Static" -- logic levels remain valid so long as power is supplied
- NOR and NAND gates



(a)



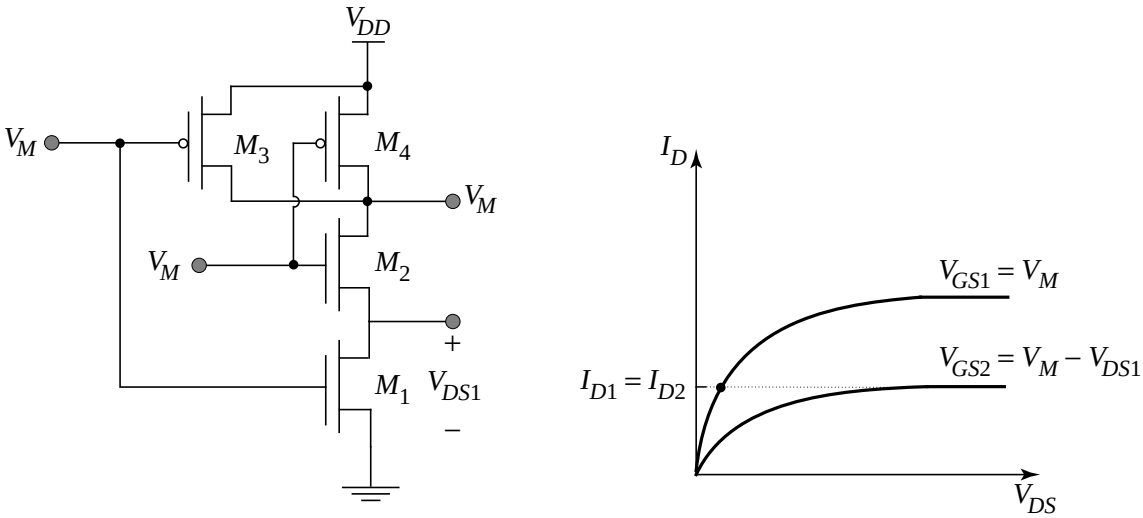
(b)

CMOS NAND Gate

- Qualitative description

  Find transfer curve for case where $V_A = V_B$ and both transition from 0 to 5 V

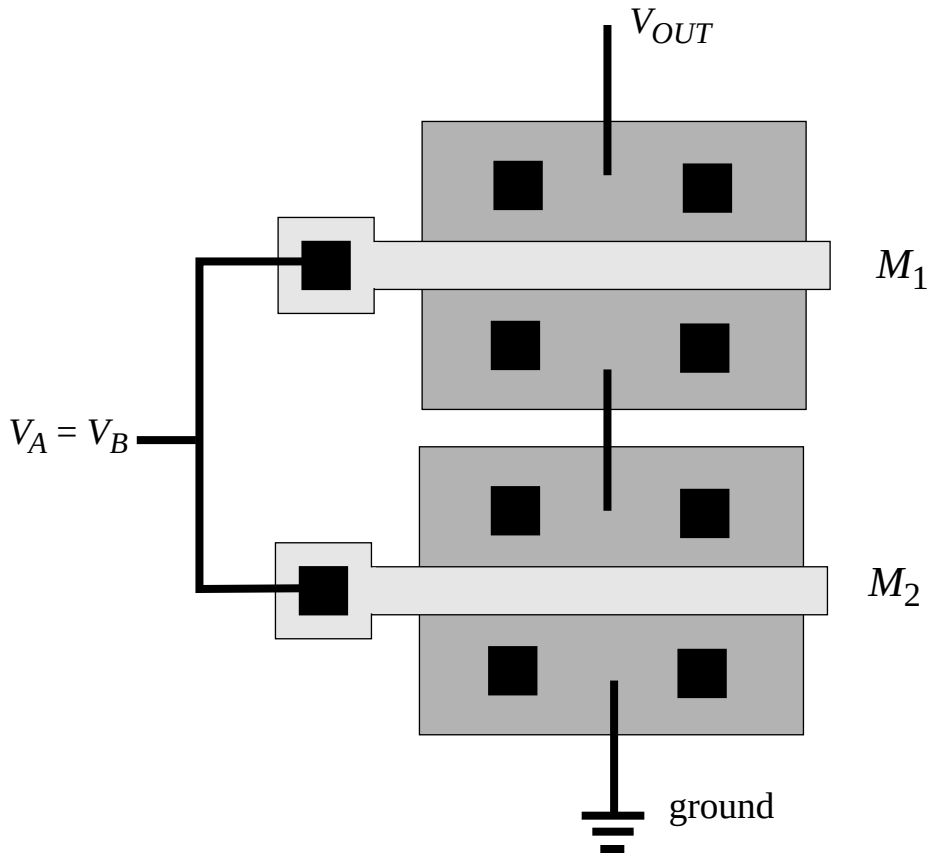
- Transistors M_1 and M_2 are in series and have the same current; however, they do *not* have the same gate-source bias



(a) (b)

MOSFETs in Series

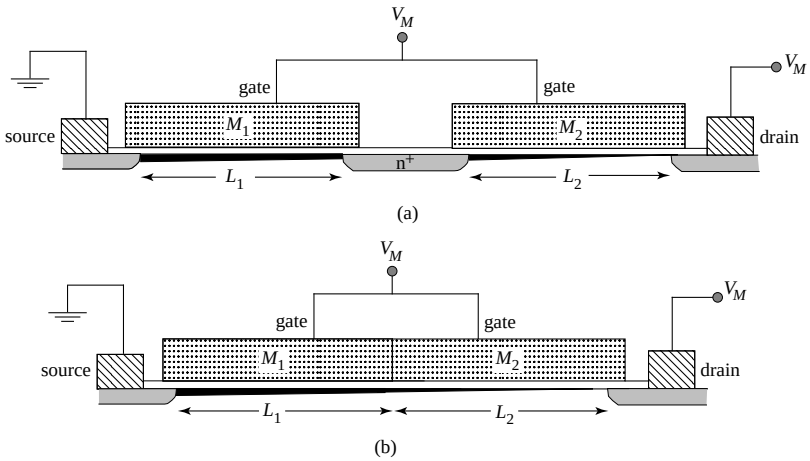
- Transistors M_1 and M_2 are "in series" with the same gate voltage, for the case where the inputs are tied together ($A = B$)



drain current is the same through each device ... what is the effective value of k_P?

MOSFETs in Series (Cont.)

- At $V_A = V_B = V_M$, the cross section through M_1 - M_2 is:



- Transistor M_1 is in triode and M_2 is saturated. From the cross section, the drain of M_1/ source of M_2 can be eliminated without affecting anything --> the two MOSFETs can be merged into a composite transistor with $L_1 + L_2 = 2\ L_n$
- Solving for V_M for the case where $V_A = V_B$ (note that the two p-channel devices are in parallel and have an effective width of $W_3 + W_4 = 2\ W_p$

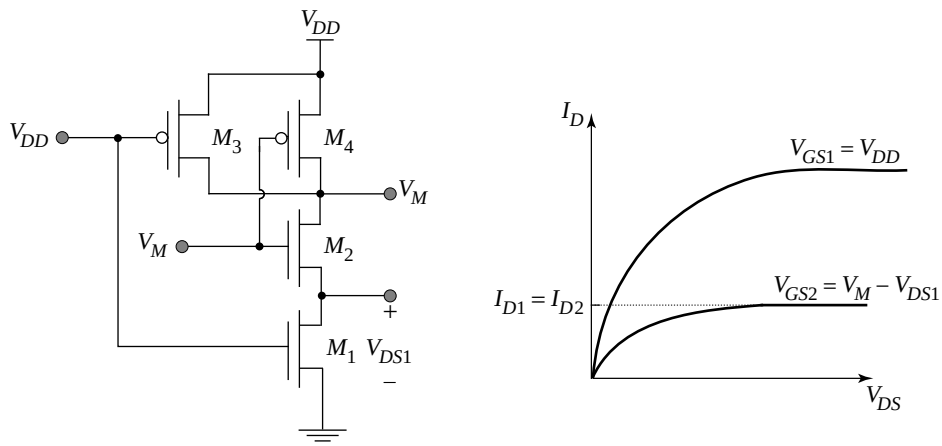
$$V_M = \frac{V_{Tn} + \sqrt{\frac{2k_p}{k_n/2}}(V_{DD} + V_{Tp})}{1 + \sqrt{\frac{2k_p}{k_n/2}}} = \frac{V_{Tn} + 2\sqrt{\frac{k_p}{k_n}}(V_{DD} + V_{Tp})}{1 + 2\sqrt{\frac{k_p}{k_n}}}$$

where $k_n = \mu_n C_{ox}\ (W_n/L_n)$ and $k_p = \mu_p C_{ox}\ (W_p/L_p)$

We could optimize $V_M = V_{DD}/2$, but there is another switching condition to consider

CMOS Static NAND Gate

- Second switching condition: $V_A = V_{DD}$ and V_B switches from 0 to V_{DD}



(a) (b)

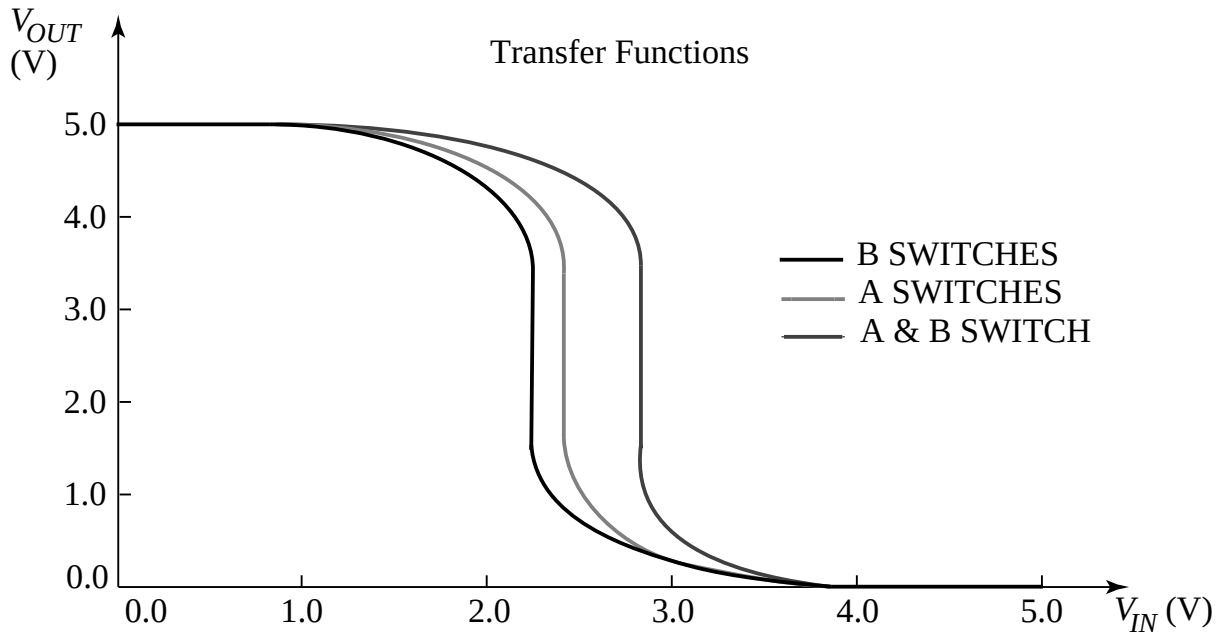
At $V_B = V_M$, the current through M_1 and M_2 is higher than when $V_A = V_B$ since the gate voltage on M_1 is now V_{DD} and its V_{DS1} must be smaller --> V_{GS2} is larger. Effective k_n is increased.

At $V_B = V_M$, only M_4 is conducting current --> only half the current as for previous switching condition. Effective k_p is that of device M_4

- Result: V_M is 0.3-0.5 V lower than for $V_A = V_B$ switching condition, for typical dimensions

NAND Gate Transfer Functions

- SPICE is useful to solve for the transfer functions under the various switching conditions (see Ex. 5.7). Note that the backgate effect means that the curves when V_A switches and when V_B switches are not identical.



- Results: setting $k_n = 2\ k_p$ results in V_M approximately $V_{DD}/2$.

CMOS NAND Gate Transient Analysis

- Worst-case situation for low-to-high transition: only one of the p-channel transistors is switching (say M_4):

$$-I_{Dp} = -I_{D4} = \frac{k_p}{2}(V_{DD} + V_{Tp})^2$$

- For high-to-low transition, consider M_1 and M_2 in series with effective length at $2L_n$ (worst-case since current is lowest with $V_A = V_B$)

$$I_{Dn} = I_{D_1} = I_{D2} = \mu_n C_{ox}[W_n/2(2L_n)](V_{DD} - V_{Tn})^2 = \frac{k_n}{4}(V_{DD} - V_{Tn})^2$$

- For equal propagation delays, we require $I_{Dn} = -I_{Dp}$

$$\frac{k_n}{4} = \frac{k_p}{2} \text{ --> } k_n = 2k_p$$

The factor of 2 mobility difference between the p and n channels indicates that

$$(W/L)_n = (W/L)_p \text{ (2 input NAND gate)}$$

- For an M-input NAND gate, we find that

$$(W/L)_n = (M/2)\,(W/L)_p$$

Note: NOR gates suffer from a factor of $2M$ between the n- and p-channel ratios which makes them unattractive for large fan-in gates