# VLSI Design Issues

Scaling/Moore's Law has limits due to the physics of material.

- Now L (L=20nm??) affects tx delays (speed), ***noise***, heat (power consumption)
- Scaling increases density of txs and requires "more" interconnect (highways & buses)-more delays (lowering speed) and heat.

Possible Solutions:

- New fabrication solutions/material. E.g., Interconnect layers, new material (copper & low k-material)
- **Improve physical Designs at the transistor level. Create better cell libraries (min Power, min-delay, max speed)**
- Exploit transister analog/physics characteristics
- Invent new transistors
- Invent new architectures

# CMOS Inverter: DC Analysis

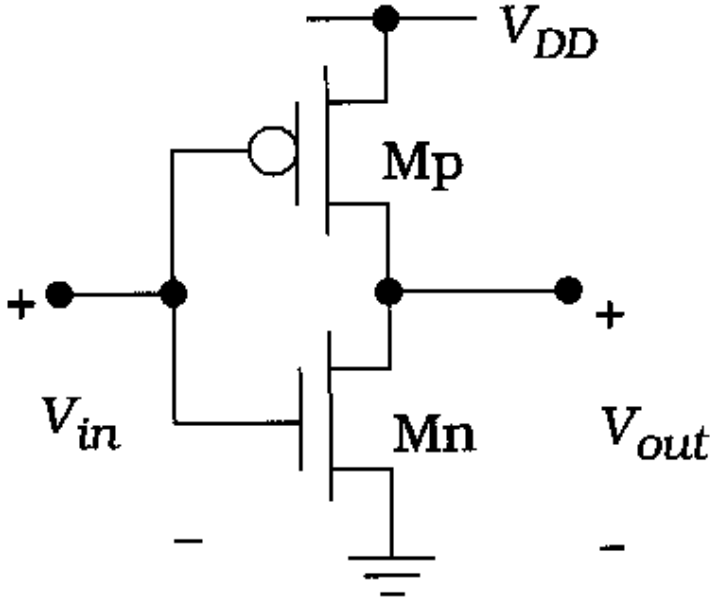
- Analyze DC Characteristics of CMOS Gates by studying an Inverter

pFET: $V_{Tp} < 0$

$$\beta_p = k'_p \left(\frac{W}{L}\right)_p$$

- DC Analysis
  - DC value of a signal in static conditions

- DC Analysis of CMOS Inverter
  - Vin, input voltage
  - Vout, output voltage
  - single power supply, VDD
  - Ground reference
  - find Vout = f(Vin)



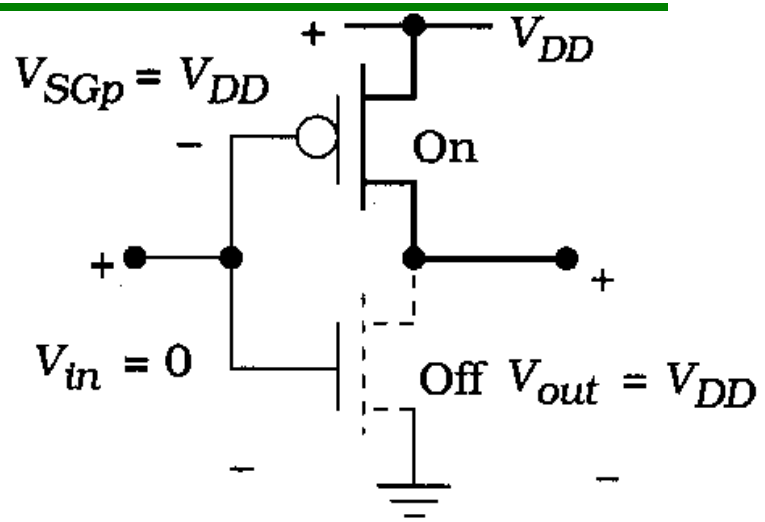

- Voltage Transfer Characteristic (VTC)
  - plot of Vout as a function of Vin
  - vary Vin from 0 to VDD (and in reverse!)
  - find Vout at each value of Vin

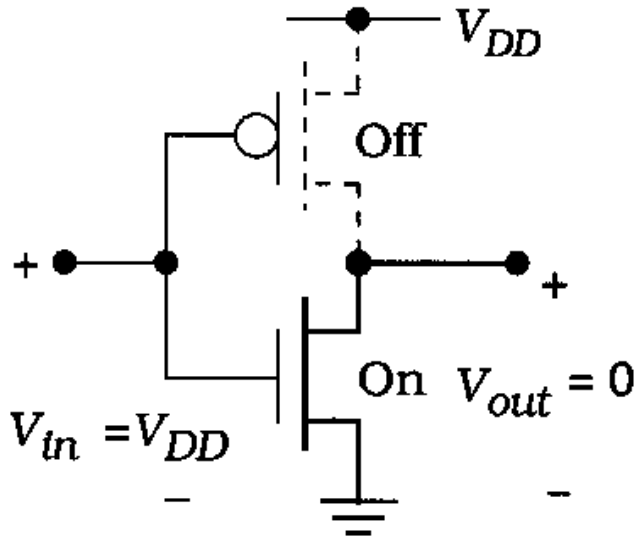
nFET: $V_{Tn} > 0$

$$\beta_n = k'_n \left(\frac{W}{L}\right)_n$$

# Inverter Voltage Transfer Characteristics

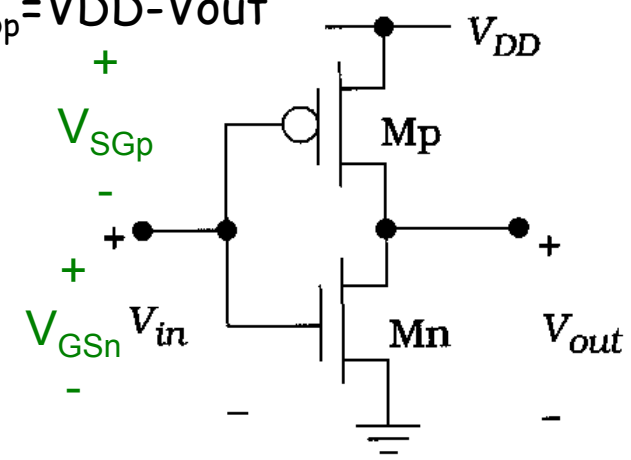
- Output High Voltage, $V_{OH}$
  - maximum output voltage
    - occurs when input is low (Vin = 0V)
    - pMOS is ON, nMOS is OFF
    - pMOS pulls Vout to VDD
  - $V_{OH}$ = VDD
- Output Low Voltage, $V_{OL}$
  - minimum output voltage
    - occurs when input is high (Vin = VDD)
    - pMOS is OFF, nMOS is ON
    - nMOS pulls Vout to Ground
  - $V_{OL}$ = 0 V
- Logic Swing
  - Max swing of output signal
    - $V_L = V_{OH} - V_{OL}$
    - $V_L$ = VDD



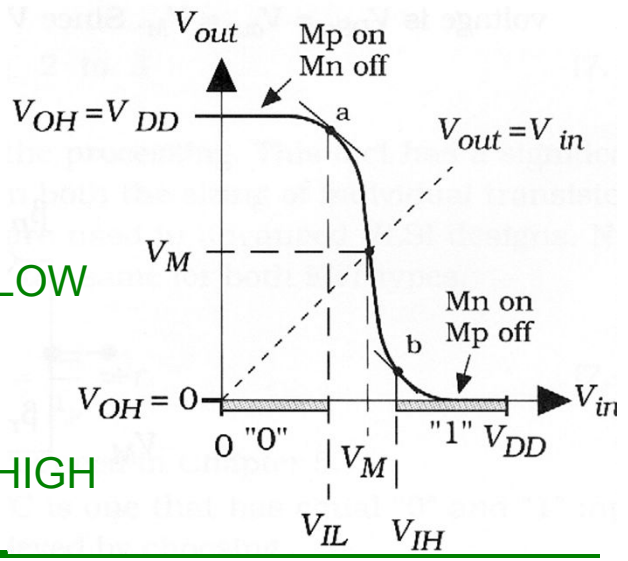



# Inverter Voltage Transfer Characteristics

- Gate Voltage, f(Vin)
  - $V_{GSn}$=Vin, $V_{SGp}$=VDD-Vin
- Drain Voltage, f(Vout)
  - $-V_{DSn}$=Vout, $V_{SDp}$=VDD-Vout




- Transition Region (between $V_{OH}$ and $V_{OL}$)
  - Vin low
    - Vin < Vtn
      - Mn in Cutoff, OFF
      - Mp in Triode, Vout pulled to VDD
    - Vin > Vtn < ~Vout
      - Mn in Saturation, strong current
      - Mp in Triode, $V_{SG}$ & current reducing
      - Vout decreases via current through Mn
  - Vin = Vout (mid point) ≈ ½ VDD
    - Mn and Mp both in Saturation
    - maximum current at Vin = Vout
  - Vin high
    - Vin > ~Vout, Vin < VDD - |Vtp|
      - Mn in Triode, Mp in Saturation
    - Vin > VDD - |Vtp|
      - Mn in Triode, Mp in Cutoff

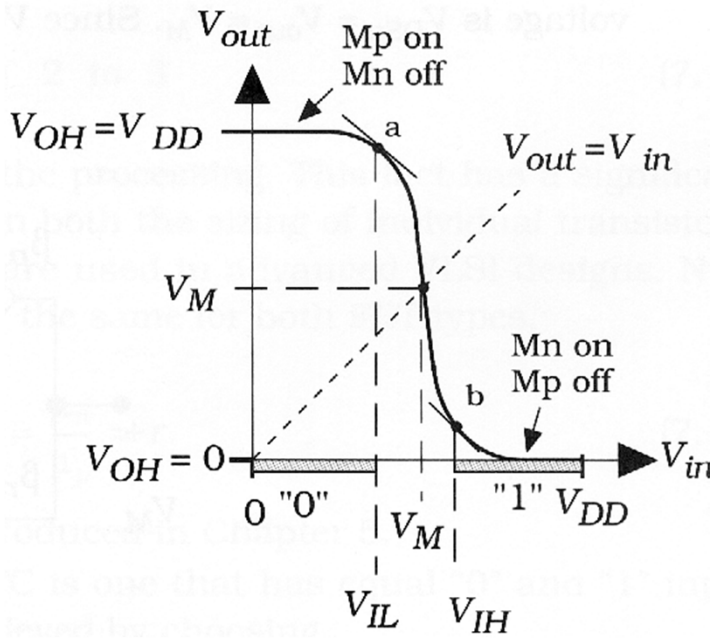



Vin < $V_{IL}$
input logic LOW

Vin > $V_{IH}$
input logic HIGH

# Noise Margin

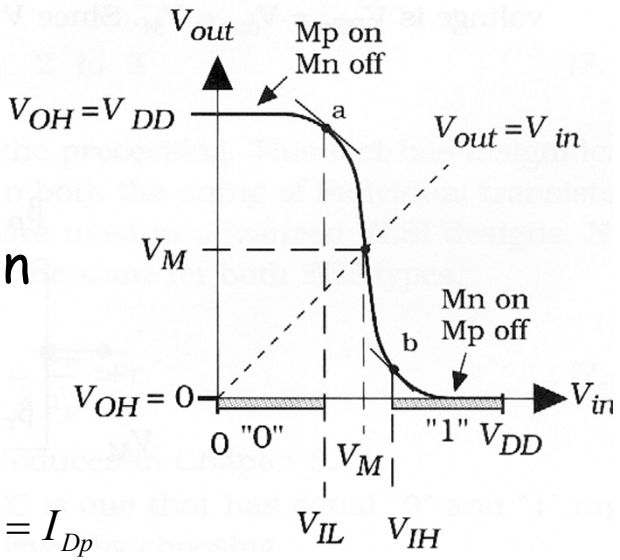
- Input Low Voltage, $V_{IL}$
  - Vin such that Vin < $V_{IL}$ = logic 0
  - point 'a' on the plot
    - where slope, $\frac{\partial Vin}{\partial Vout} = -1$
- Input High Voltage, $V_{IH}$
  - Vin such that Vin > $V_{IH}$ = logic 1
  - point 'b' on the plot
    - where slope =-1




- Voltage Noise Margins
  - measure of how stable inputs are with respect to signal interference
  - $VNM_H = V_{OH} - V_{IH}$ $= VDD - V_{IH}$
  - $VNM_L = V_{IL} - V_{OL}$ $= V_{IL}$
  - desire large $VNM_H$ and $VNM_L$ for best noise immunity

# Switching Threshold

- Switching threshold = point on VTC where Vout = Vin
  - also called midpoint voltage, $V_M$
  - here, Vin = Vout = $V_M$
- Calculating $V_M$
  - at $V_M$, both nMOS and pMOS in Saturation
  - in an inverter, $I_{Dn} = I_{Dp}$, always!
  - solve equation for $V_M$




$$I_{Dn} = \frac{\mu_n C_{OX}}{2}\frac{W}{L}(V_{GSn} - V_{tn})^2 = \frac{\beta_n}{2}(V_{GSn} - V_{tn})^2 = \frac{\beta_p}{2}(V_{SGp} - |V_{tp}|)^2 = I_{Dp}$$

  - express in terms of $V_M$

$$\frac{\beta_n}{2}(V_M - V_{tn})^2 = \frac{\beta_p}{2}(V_{DD} - V_M - |V_{tp}|)^2 \quad \Rightarrow \quad \sqrt{\frac{\beta_n}{\beta_p}}(V_M - V_{tn}) = V_{DD} - V_M - |V_{tp}|$$

  - solve for $V_M$

$$V_M = \frac{VDD - |V_{tp}| + V_{tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}}$$

# Effect of Transistor Size on VTC

- Recall

$$\beta_n = k'_n \frac{W}{L} \qquad \frac{\beta_n}{\beta_p} = \frac{k'_n \left(\frac{W}{L}\right)_n}{k'_p \left(\frac{W}{L}\right)_p}$$

$$V_M = \frac{VDD - |V_{tp}| + V_{tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}}$$

- If nMOS and pMOS are same size
  - (W/L)n = (W/L)p
  - Coxn = Coxp (always)

$$\frac{\beta_n}{\beta_p} = \frac{\mu_n C_{oxn}\left(\frac{W}{L}\right)_n}{\mu_p C_{oxp}\left(\frac{W}{L}\right)_p} = \frac{\mu_n}{\mu_p} \cong 2 or 3$$

- If $\frac{\left(\frac{W}{L}\right)_p}{\left(\frac{W}{L}\right)_n} = \frac{\mu_n}{\mu_p}, then \frac{\beta_n}{\beta_p} = 1$

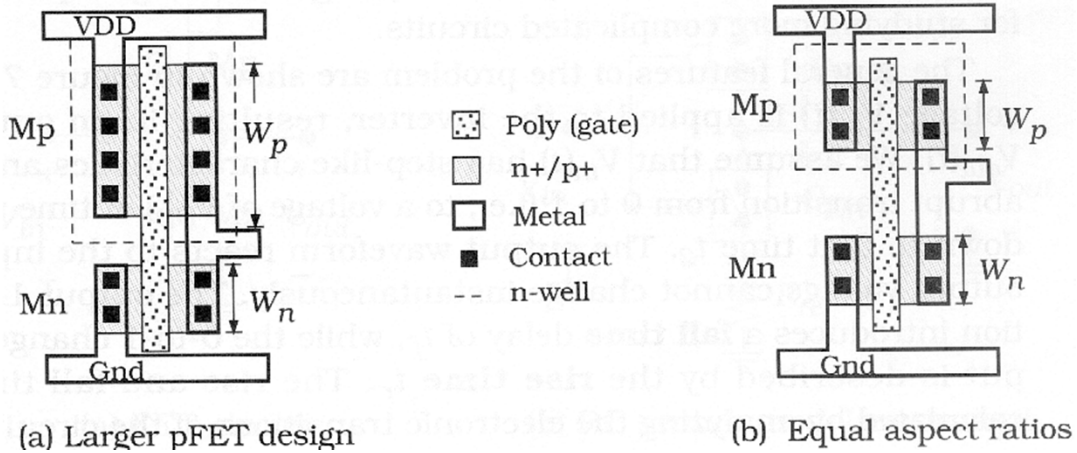
since L normally min. size for all tx,
can get betas equal by making Wp larger than Wn

- Effect on switching threshold
  - if $\beta_n \approx \beta_p$ and Vtn = |Vtp|, $V_M$ = VDD/2, <u>exactly in the middle</u>

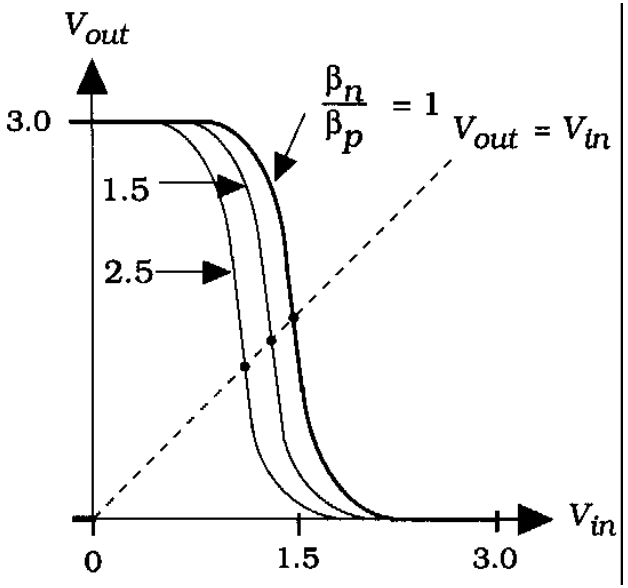
- Effect on noise margin
  - if $\beta_n \approx \beta_p$, $V_{IH}$ and $V_{IL}$ both close to $V_M$ and <u>noise margin is good</u>

# Example

- Given
  - k'n = 140uA/V², Vtn = 0.7V, VDD = 3V
  - k'p = 60uA/V², Vtp = -0.7V
- Find
  - a) tx size ratio so that $V_M$= 1.5V
  - b) $V_M$ if tx are same size



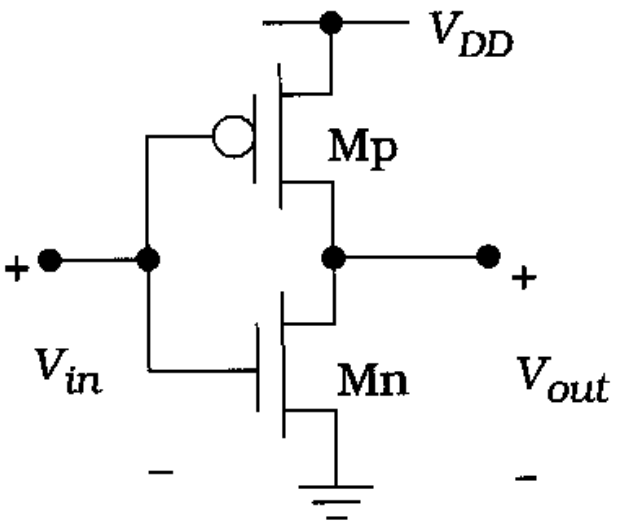

(a) Larger pFET design (b) Equal aspect ratios

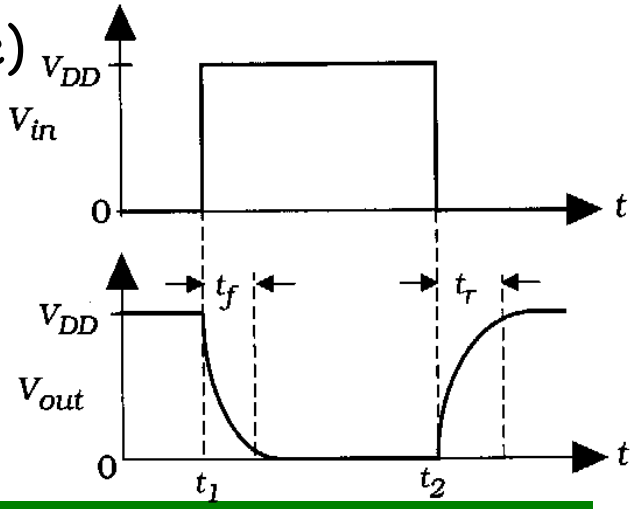



transition pushed lower as beta ratio increases

# CMOS Inverter: Transient Analysis

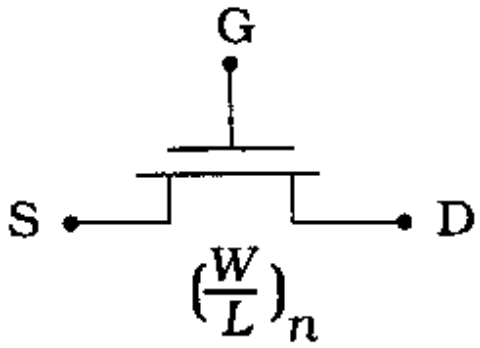
- Analyze Transient Characteristics of CMOS Gates by studying an Inverter
- Transient Analysis
  - signal value as a function of time
- Transient Analysis of CMOS Inverter
  - Vin(t), input voltage, function of time
  - Vout(t), output voltage, function of time
  - VDD and Ground, DC (not function of time)
  - find Vout(t) = f(Vin(t))
- Transient Parameters
  - output signal rise and fall time
  - propagation delay

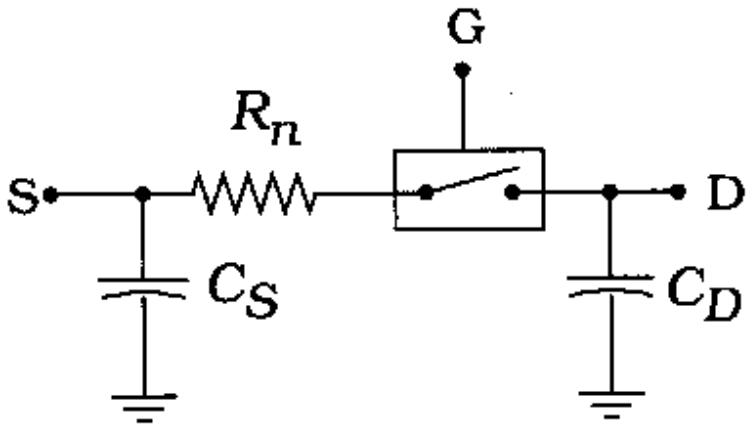


# Transient Response

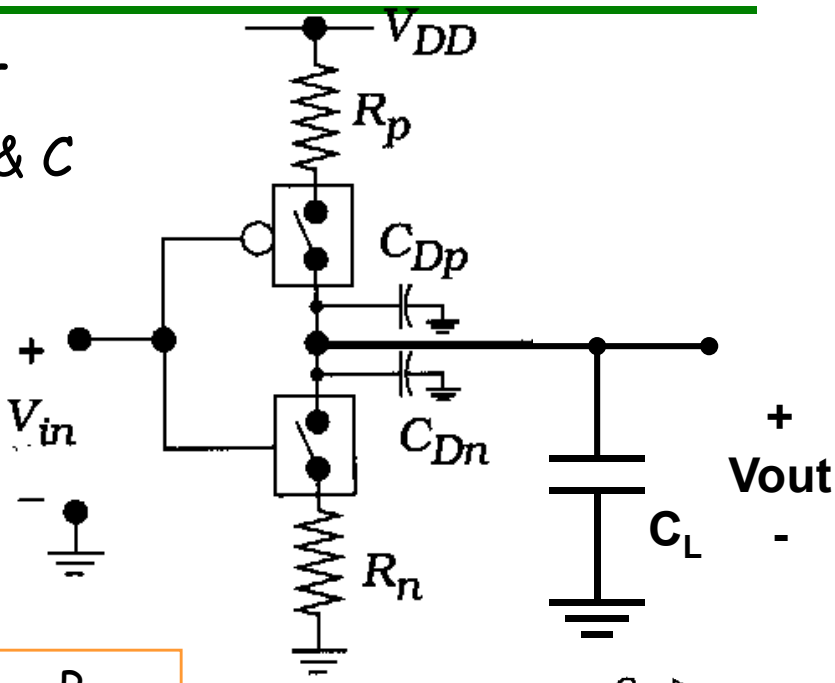
- Recall: the RC nMOS Transistor Model

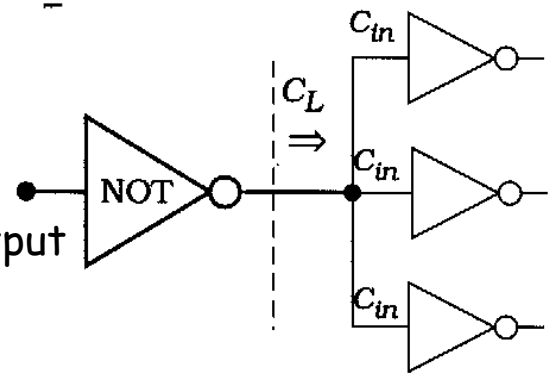


# Transient Response

- Response to step change in input
  - delays in output due to parasitic R & C
- Inverter RC Model
  - Resistances (linear model)
    - $R_n = 1/[\beta_n(V_{DD}-V_{tn})]$
    - $R_p = 1/[\beta_p(V_{DD}-|V_{tp}|)]$
  - Output Cap. (only output is important)
    - $C_{Dn}$ (nMOS drain capacitance)
      - $C_{Dn} = \frac{1}{2} C_{ox} W_n L + C_j A_{Dnbot} + C_{jsw} P_{Dnsw}$
    - $C_{Dp}$ (pMOS drain capacitance)
      - $C_{Dp} = \frac{1}{2} C_{ox} W_p L + C_j A_{Dpbot} + C_{jsw} P_{Dpsw}$
    - Load capacitance, due to gates attached at the output
      - $C_L = 3\,Cin = 3\,(C_{Gn} + C_{Gp})$, 3 is a "typical" load
    - Total Output Capacitance
      - $Cout = C_{Dn} + C_{Dp} + C_L$




term "**fan-out**" describes
# gates attached at output

# Fall Time

- Fall Time, $t_f$
  - time for output to **fall** from '1' to '0'
  - derivation:

$$i = -C_{out}\frac{\partial V_{out}}{\partial t} = \frac{V_{out}}{R_n}$$

    - initial condition, Vout(0) = VDD
    - solution

$$Vout(t) = V_{DD}e^{-t/\tau_n}$$

time constant $\tau_n = R_n C_{out}$

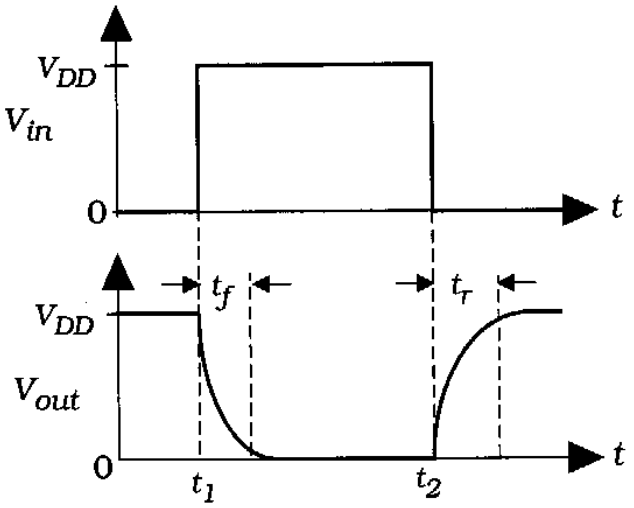
$$t = \tau_n \ln\left(\frac{V_{DD}}{Vout}\right)$$

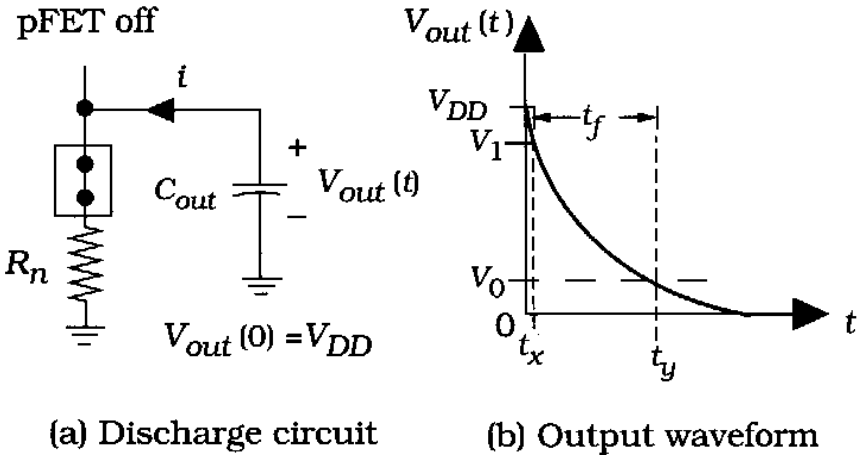
  - definition
    - $t_f$ is time to fall from

90% value $[V_1, t_x]$ to 10% value $[V_0, t_y]$

$$t = \tau_n\left[\ln\left(\frac{V_{DD}}{0.1V_{DD}}\right) - \ln\left(\frac{V_{DD}}{0.9V_{DD}}\right)\right]$$

    - $t_f = 2.2\,\tau_n$




(a) Discharge circuit (b) Output waveform

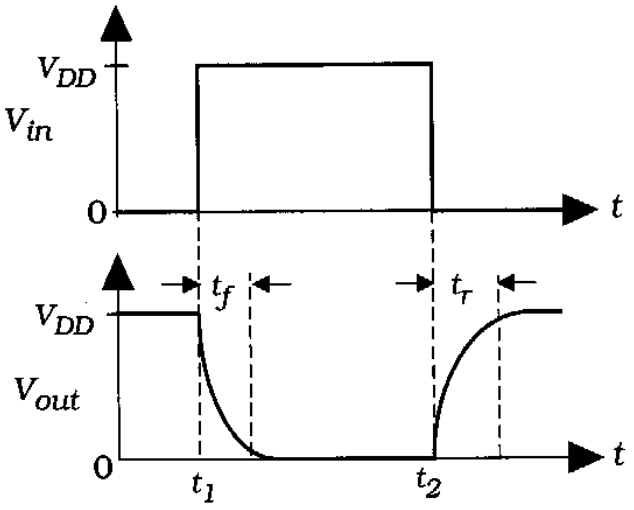
# Rise Time

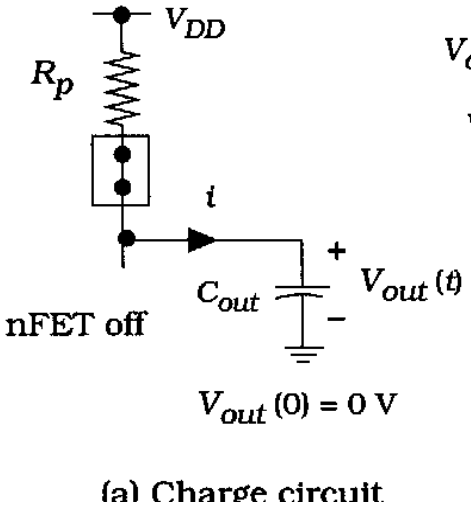
- Rise Time, $t_r$
  - time for output to **rise** from '0' to '1'
  - derivation: $$i = C_{out}\frac{\partial V_{out}}{\partial t} = \frac{V_{DD} - V_{out}}{R_p}$$
    - initial condition, Vout(0) = 0V
    - solution

$$Vout(t) = V_{DD}\left[1 - e^{-t/\tau_p}\right]$$

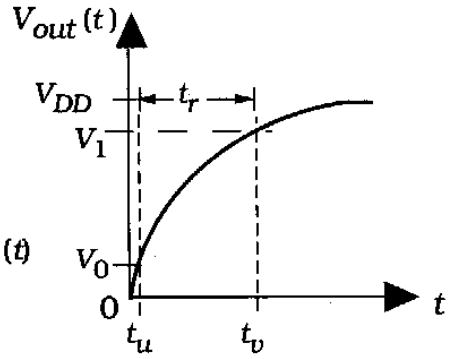
time constant $\tau_p = R_p C_{out}$

  - definition
    - $t_f$ is time to rise from 10% value $[V_0, t_u]$ to 90% value $[V_1, t_v]$
    - $t_r = 2.2\,\tau_p$
- Maximum Signal Frequency
  - $f_{max} = 1/(t_r + t_f)$
    - faster than this and the output can't settle



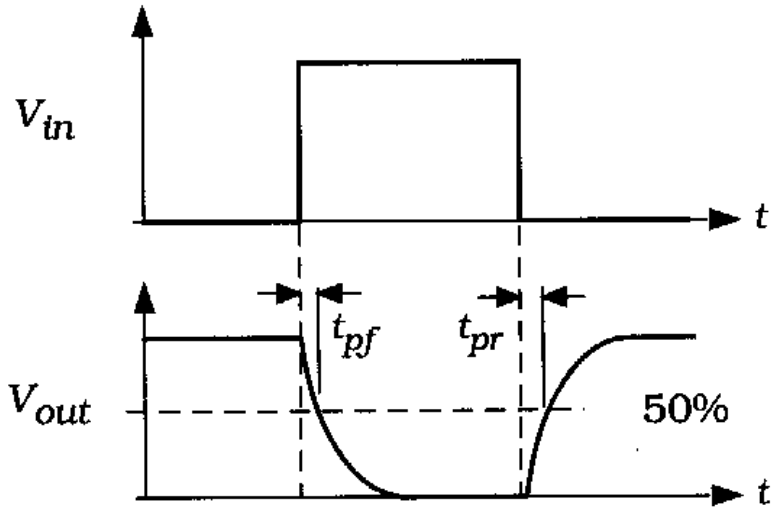

(a) Charge circuit




(b) Output waveform

# Propagation Delay

- Propagation Delay, $t_p$
  - measures speed of output reaction to input change
  - $t_p = \frac{1}{2}(t_{pf} + t_{pr})$
- Fall propagation delay, $t_{pf}$
  - time for output to fall by 50%
    - reference to input switch
- Rise propagation delay, $t_{pr}$
  - time for output to rise by 50%
    - reference to input switch




- Ideal expression (if input is step change)
  - $t_{pf} = \ln(2)\,\tau_n$
  - $t_{pr} = \ln(2)\,\tau_p$
- Total Propagation Delay
  - $t_p = 0.35(\tau_n + \tau_p)$

Propagation delay measurement:
- from time input reaches 50% value
- to time output reaches 50% value

Add rise and fall propagation delays for total value

# Switching Speed -Resistance

- Rise & Fall Time
  - $t_f = 2.2\,\tau_n$, $t_r = 2.2\,\tau_p$,
- Propagation Delay
  - $t_p = 0.35(\tau_n + \tau_p)$
- In General
  - delay $\propto \tau_n + \tau_p$
  - $\tau_n + \tau_p = C_{out}(R_n + R_p)$
- Define delay in terms of design parameters
  - $R_n + R_p = \dfrac{(V_{DD}-V_t)(\beta_n+\beta_p)}{\beta_n \beta_p (V_{DD}-V_t)^2}$
  - $R_n + R_p = \dfrac{\beta_n + \beta_p}{\beta_n \beta_p (V_{DD}-V_t)}$
    - if $V_t = V_{tn} = |V_{tp}|$

$\tau_n = R_n C_{out}$ $\quad$ $\tau_p = R_p C_{out}$

$R_n = 1/[\beta_n(V_{DD}-V_{tn})]$ $\quad$ $\beta = \mu C_{ox}(W/L)$

$R_p = 1/[\beta_p(V_{DD}-|V_{tp}|)]$

$C_{out} = C_{Dn} + C_{Dp} + C_L$

Beta Matched if $\beta_n=\beta_p=\beta$,

$$R_n+R_p = \frac{2}{\beta(V_{DD}-V_t)} = \frac{2L}{\mu C_{ox} W (V_{DD}-V_t)}$$

Width Matched if $W_n=W_p=W$, and $L=L_n=L_p$

$$R_n+R_p = \frac{L(\mu_n+\mu_p)}{(\mu_n \mu_p) C_{ox} W (V_{DD}-V_t)}$$

To decrease R's, ⇓L, ⇑W, ⇑VDD, ( ⇑$\mu_p$, ⇑Cox )

# Switching Speed -Capacitance

- From Resistance we have
  - ⇓L, ⇑W, ⇑VDD, ( ⇑$\mu_p$, ⇑Cox )
  - but ⇑ VDD increases power
  - ⇑ W increases Cout
- Cout
  - Cout = $\frac{1}{2}$ Cox L ($W_n$+$W_p$) + $C_j$ 2L ($W_n$+$W_p$) + 3 Cox L ($W_n$+$W_p$)
    - assuming junction area ~W•2L
    - neglecting sidewall capacitance
  - Cout ≈ L ($W_n$+$W_p$) [$3\frac{1}{2}$ Cox +2 $C_j$]
  - Cout ∝ L ($W_n$+$W_p$)

  To decrease Cout, ⇓L, ⇓W, (⇓Cj, ⇓Cox )
- Delay ∝ Cout(Rn+Rp) ∝ L W $\frac{L}{W\ VDD} = \frac{L^2}{VDD}$

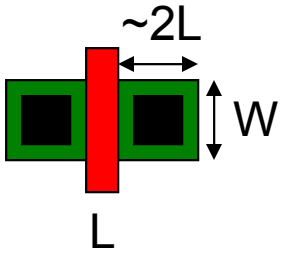
$$C_{out} = C_{Dn} + C_{Dp} + C_L$$

if $L=L_n=L_p$ estimate

$$C_L = 3\,(C_{Gn} + C_{Gp}) = 3\,Cox\,(W_nL+W_pL)$$

$$C_{Dn} = \tfrac{1}{2}\,Cox\,W_n\,L + C_j\,A_{Dnbot} + C_{jsw}\,P_{Dnsw}$$

$$C_{Dp} = \tfrac{1}{2}\,Cox\,W_p\,L + C_j\,A_{Dpbot} + C_{jsw}\,P_{Dpsw}$$



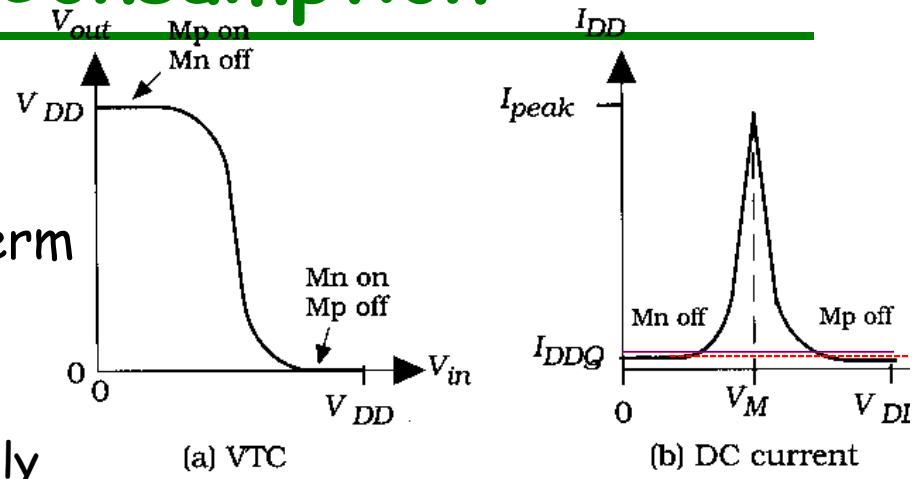

Decreasing L (reducing feature size) is best way to improve speed!

# Switching Speed -Local Modification

- Previous analysis applies to the overall design
  - shows that reducing feature size is critical for higher speed
  - general result useful for creating cell libraries
- How do you improve speed within a specific gate?
  - increasing W in one gate will not increase $C_G$ of the load gates
    - Cout = $C_{Dn} + C_{Dp} + C_L$
    - increasing W in one logic gate will increase $C_{Dn/p}$ but not $C_L$
      - $C_L$ depends on the size of the tx gates at the output
      - as long as they keep minimum W, $C_L$ will be constant
  - thus, increasing W is a good way to improve the speed within a local point
  - But, increasing W increases chip area needed, which is bad
    - fast circuits need more chip area (chip "real estate")
- Increasing VDD is not a good choice because it increases power consumption

# CMOS Power Consumption

- $P = P_{DC} + P_{dyn}$
  - $P_{DC}$: DC (static) term
  - $P_{dyn}$: dynamic (signal changing) term
- $P_{DC}$
  - $P = I_{DD}\ V_{DD}$
    - $I_{DD}$ DC current from power supply
    - ideally, $I_{DD} = 0$ in CMOS: ideally only current during switching action
    - leakage currents cause $I_{DD} > 0$, define **quiescent** leakage current, $I_{DDQ}$ (due largely to leakage at substrate junctions)
  - $P_{DC} = I_{DDQ}\ V_{DD}$
- Pdyn, power required to switch the state of a gate
  - charge transferred during transition, Qe = Cout VDD
  - assume each gate must transfer this charge 1x/clock cycle
  - Paverage = $V_{DD}$ Qe f = Cout $V_{DD}^2$ f, f = frequency of signal change
- Total Power, $P = I_{DDQ}\ V_{DD} + \text{Cout}\ V_{DD}^2\ f$

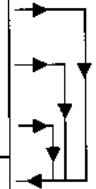



Power increases with Cout and frequency, and **strongly** with VDD (second order).

# Multi-Input Gate Signal Transitions

- In multi-input gates multiple signal transitions produce output changes
- What signal transitions need to be analyzed?
  - for a general N-input gate with $M_0$ low output states and $M_1$ high output states
    - # high-to-low output transitions = $M_0 \cdot M_1$
    - # low-to-high output transitions = $M_1 \cdot M_0$
    - total transitions to be characterized = $2 \cdot M_0 \cdot M_1$
    - example: NAND has $M_0 = 1$, $M_1 = 3$
  - don't test/characterize cases without output transitions

| | $V_A$ | $V_B$ | $V_{out}$ |
|---|---|---|---|
| (i) | 0 | 0 | $V_{DD}$ |
| (ii) | 0 | $V_{DD}$ | $V_{DD}$ |
| (iii) | $V_{DD}$ | 0 | $V_{DD}$ |
| | $V_{DD}$ | $V_{DD}$ | 0 |

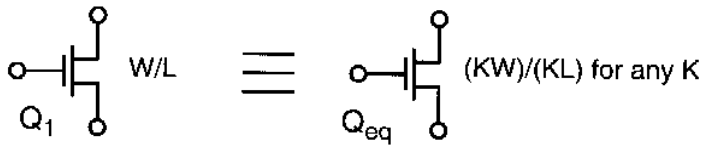


(a) Transition table

- Worst-case delay is the slowest of all possible cases
  - worst-case high-to-low
  - worst-case low-to-high
  - often different input transitions for each of these cases

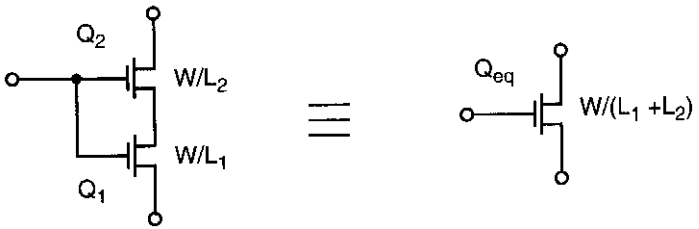
# Series/Parallel Equivalent Circuits

- Scale both W and L
  - no effective change in W/L
  - increases gate capacitance

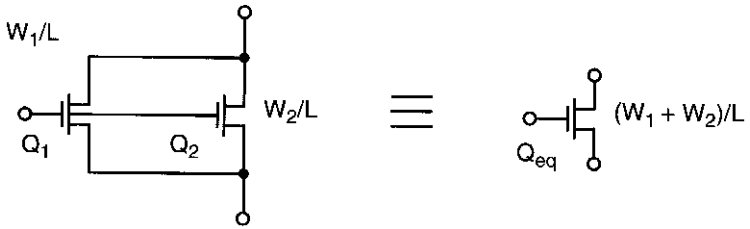



**inputs must be at same value/voltage**

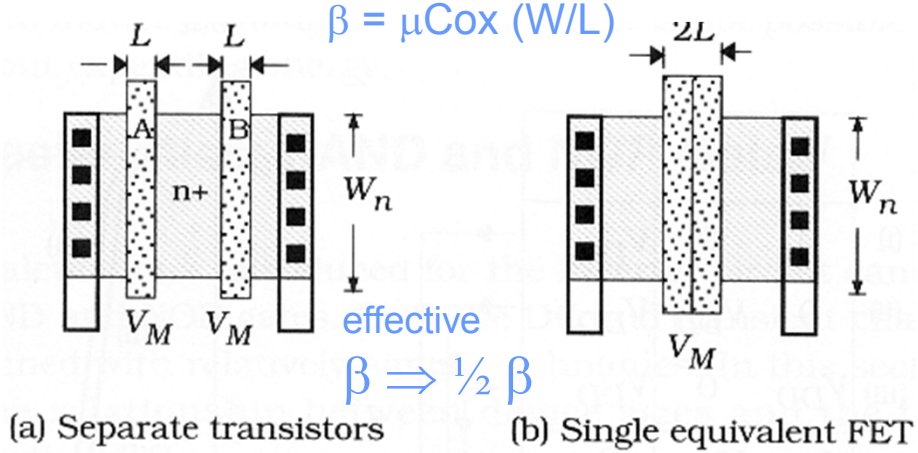
- Series Transistors
  - increases effective L

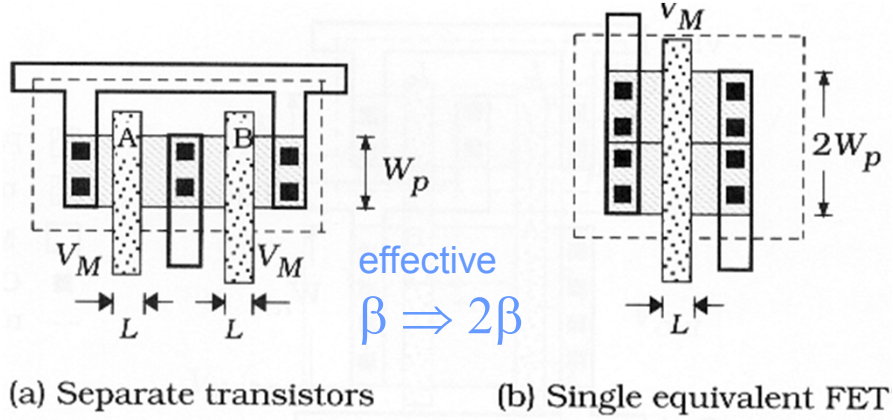



- Parallel Transistors
  - increases effective W




β = μCox (W/L)



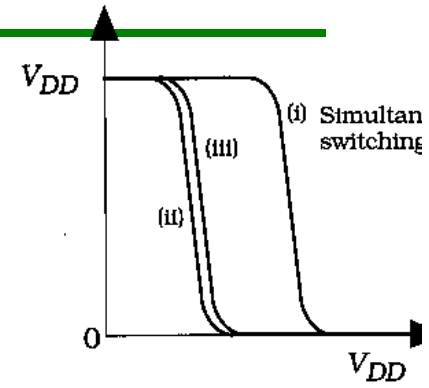



# NAND: DC Analysis

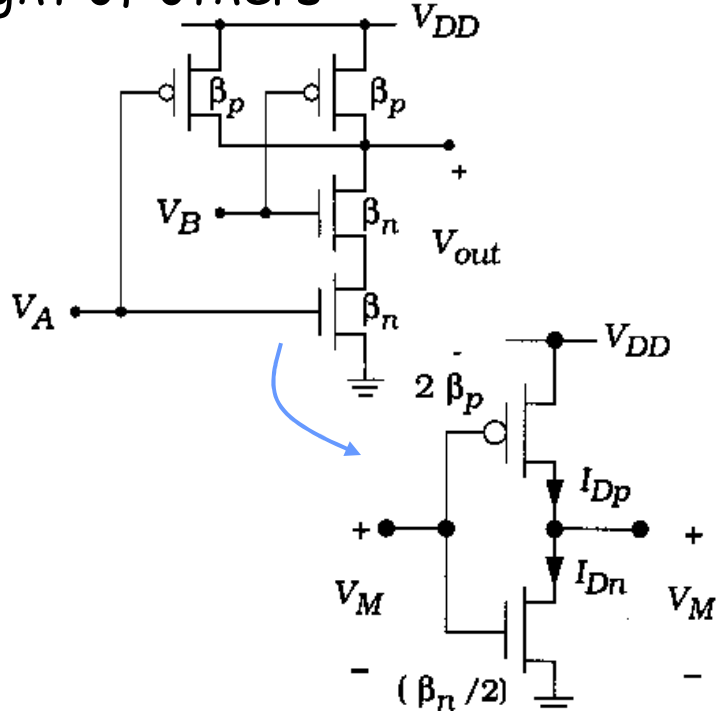
- Multiple Inputs
- Multiple Transitions
- Multiple VTCs
  - VTC varies with transition
    - transition <u>from 0,0 to 1,1</u> pushed right of others
    - why?
  - $V_M$ varies with transition
    - assume all tx have same L
    - $V_M = V_A = V_B = V_{out}$
      - can merge transistors at this point
    - if $W_{pA}=W_{pB}$ and $W_{nA}=W_{nB}$
      - series nMOS, $\beta n \Rightarrow \frac{1}{2}\beta n$
      - parallel pMOS, $\beta p \Rightarrow 2\,\beta p$
  - can now calculate the NAND $V_M$

| | $V_A$ | $V_B$ | $V_{out}$ |
|---|---|---|---|
| (i) | 0 | 0 | $V_{DD}$ |
| (ii) | 0 | $V_{DD}$ | $V_{DD}$ |
| (iii) | $V_{DD}$ | 0 | $V_{DD}$ |
| | $V_{DD}$ | $V_{DD}$ | 0 |

(a) Transition table




(b) VTC family



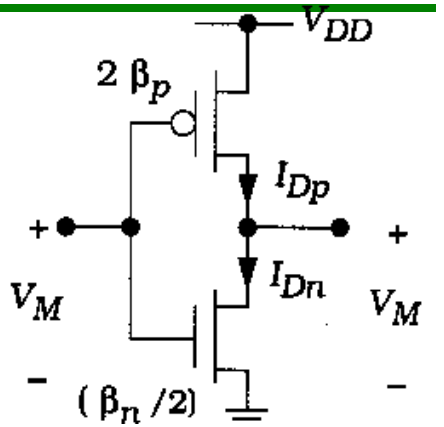
# NAND Switching Point

- Calculate VM for NAND
  - 0,0 to 1,1 transition
    - all tx change states (on, off)
    - in other transitions, only 2 change
  - $V_M = V_A = V_B = Vout$
  - set $I_{Dn} = I_{Dp}$, solve for $V_M$

$$V_M = \frac{VDD - |V_{tp}| + V_{tn}\frac{1}{2}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \frac{1}{2}\sqrt{\frac{\beta_n}{\beta_p}}}$$

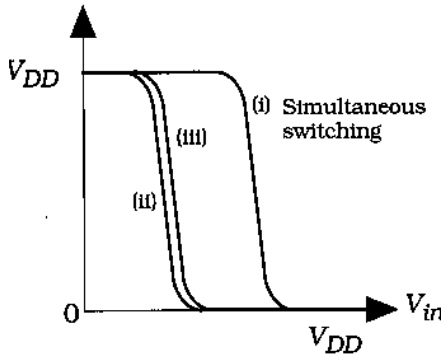
  - denominator reduced more
    - VTC shifts right
- For NAND with N inputs

$$V_M = \frac{VDD - |V_{tp}| + V_{tn}\frac{1}{N}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \frac{1}{N}\sqrt{\frac{\beta_n}{\beta_p}}}$$




series nMOS means more resistance to output **falling**,

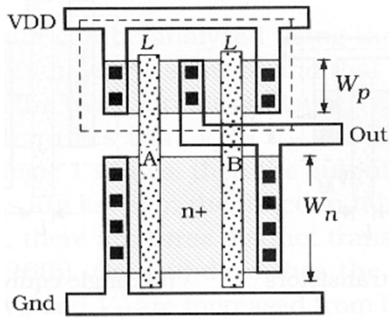
shifts VTC to right




to balance this effect and set $V_M$ to $V_{DD}/2$, can increase $\beta$ by increasing Wn

but, since $\mu_n > \mu_p$, $V_M \approx V_{DD}/2$ when Wn = Wp

(b) VTC family

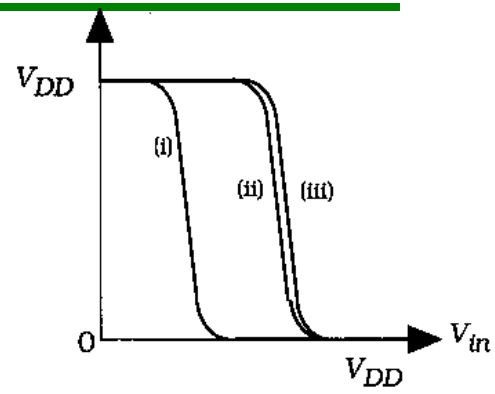


# NOR: DC Analysis

- Similar Analysis to NAND
- Critical Transition
  - 0,0 to 1,1
  - when all transistors change
- $V_M$ for NOR2 critical transition
  - if $W_{pA}=W_{pB}$ and $W_{nA}=W_{nB}$
    - parallel nMOS, $\beta n \Rightarrow 2\,\beta n$
    - series pMOS, $\beta p \Rightarrow \frac{1}{2}\,\beta p$

| | $V_A$ | $V_B$ | $V_{out}$ |
|---|---|---|---|
| | 0 | 0 | $V_{DD}$ |
| (iii) | 0 | $V_{DD}$ | 0 |
| (ii) | $V_{DD}$ | 0 | 0 |
| (i) | $V_{DD}$ | $V_{DD}$ | 0 |

(a) Transition table



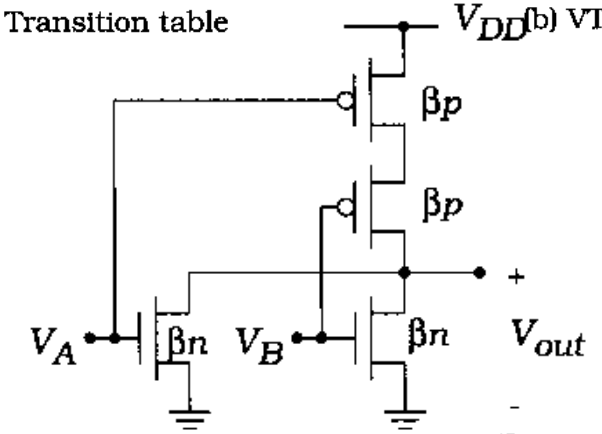

(b) VTC family

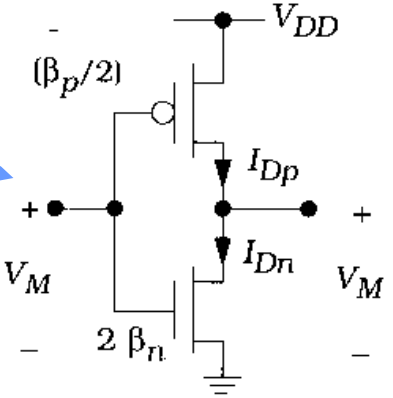
$$V_M = \frac{VDD - |V_{tp}| + 2V_{tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + 2\sqrt{\frac{\beta_n}{\beta_p}}}$$

*for NOR2*

$$V_M = \frac{VDD - |V_{tp}| + NV_{tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + N\sqrt{\frac{\beta_n}{\beta_p}}}$$

*for NOR-N*

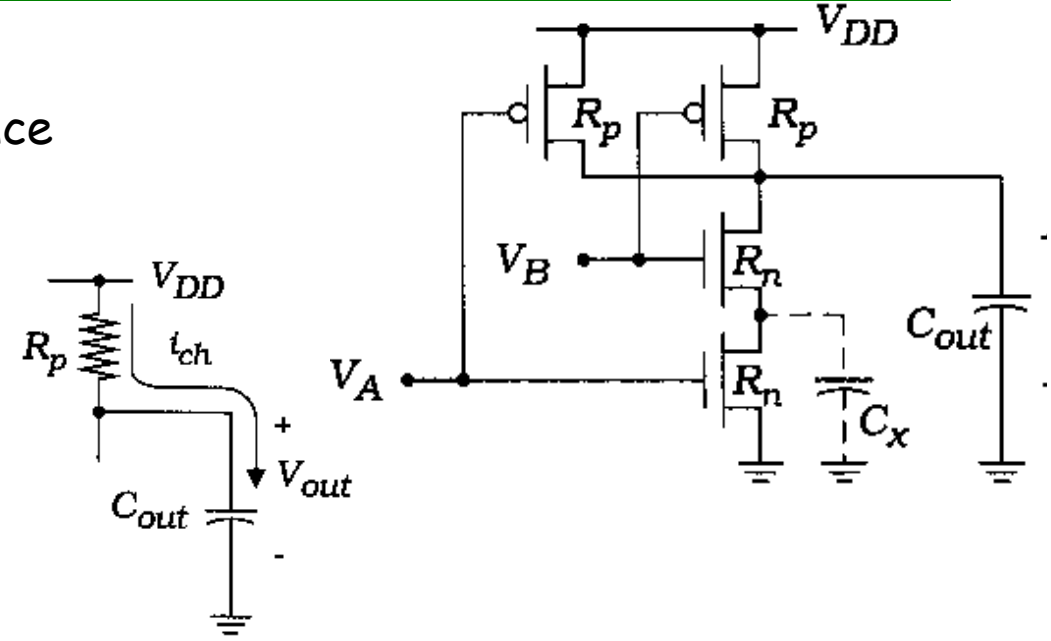


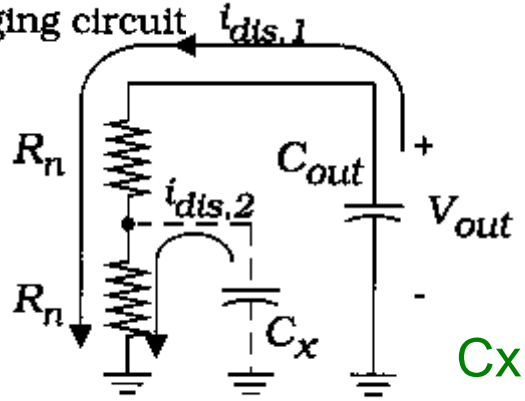

- series pMOS resistance means slower rise
- VTC shifted to the left
- to set $V_M$ to $V_{DD}/2$, increase Wp
  - this will increase βp

# NAND: Transient Analysis

- NAND RC Circuit
  - R: standard channel resistance
  - C: Cout = $C_L + C_{Dn} + 2C_{Dp}$
- Rise Time, $t_r$
  - Worst case charge circuit
    - 1 pMOS ON
  - $t_r = 2.2\ \tau_p$
    - $\tau_p = R_p$ Cout
  - best case charge circuit
    - 2 pMOS ON, Rp ⇒ Rp/2
- Fall Time, $t_f$
  - Discharge Circuit
    - 2 series nMOS, Rn ⇒ 2Rn
    - must account for internal cap, Cx
  - $t_f = 2.2\ \tau_n$
    - $\tau_n$ = Cout (2 $R_n$ ) + Cx $R_n$




(a) Charging circuit

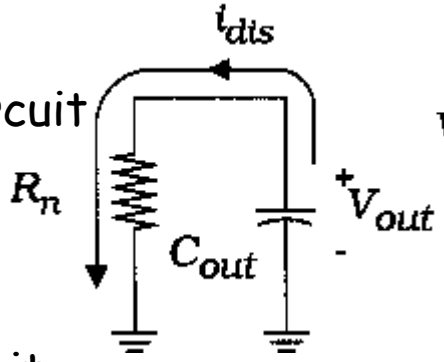


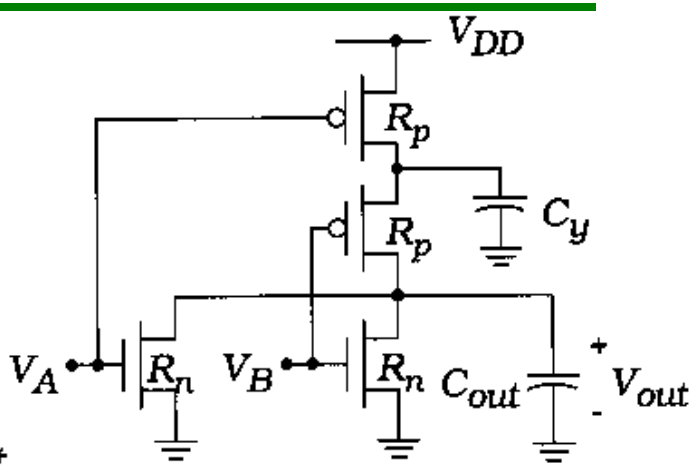

(b) Discharging circuit

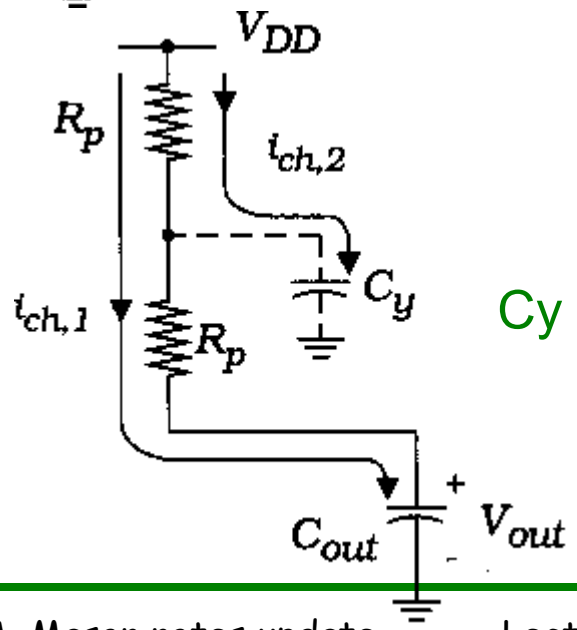
Cx = $C_{Sn} + C_{Dn}$

# NOR: Transient Analysis

- NAND RC Circuit
  - R: standard channel resistance
  - C: Cout = $C_L + 2C_{Dn} + C_{Dp}$
- Fall Time, $t_f$
  - Worst case discharge circuit
    - 1 nMOS ON
  - $t_f = 2.2\ \tau_n$
    - $\tau_n = R_n$ Cout
  - best case discharge circuit
    - 2 nMOS ON, Rn ⇒ Rn/2
- Rise Time, $t_r$
  - Charge Circuit
    - 2 series pMOS, Rp ⇒ 2Rp
    - must account for internal cap, Cy
  - $t_r = 2.2\ \tau_p$
    - $\tau_p$ = Cout (2 $R_p$ ) + Cy $R_p$



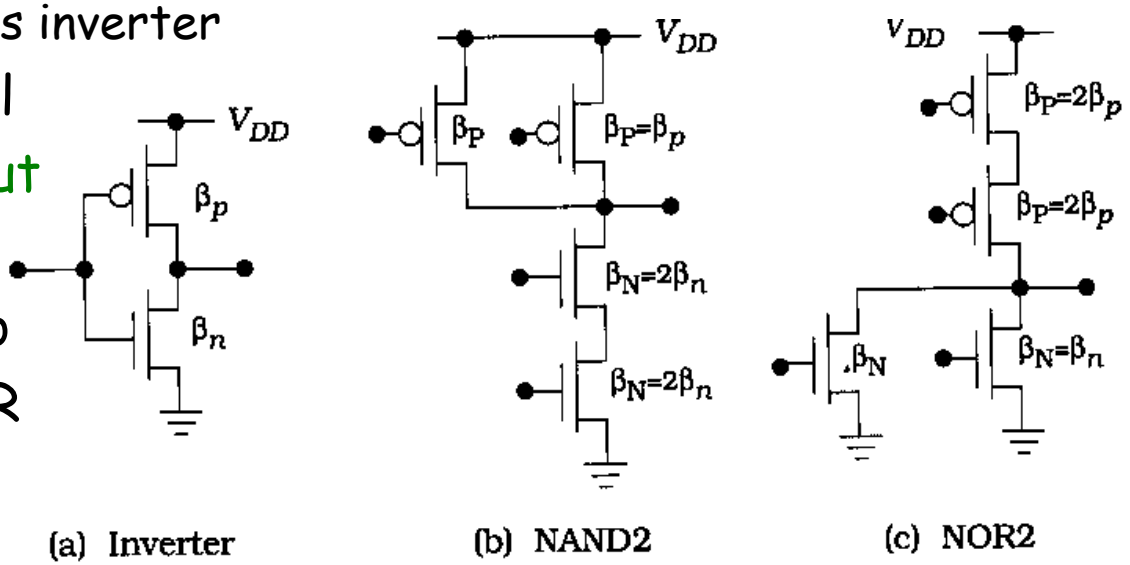




Cy = $C_{Sp} + C_{Dp}$

# NAND/NOR Performance

- Inverter: symmetry ($V_M=V_{DD}/2$), $\beta n = \beta p$
  - $(W/L)_p = \mu_n/\mu_p\ (W/L)_n$
- Match INV performance with NAND
  - pMOS, $\beta_P = \beta p$, same as inverter
  - nMOS, $\beta_N = 2\beta n$, to balance for 2 series nMOS
- Match INV performance with NOR
  - pMOS, $\beta_P = 2\ \beta p$, to balance for 2 series pMOS
  - nMOS, $\beta_N = \beta n$, same as inverter
- NAND and NOR will still be slower due to larger Cout
- This can be extended to 3, 4, … N input NAND/NOR gates

$\beta$ is adjusted by changing transistor size (width)

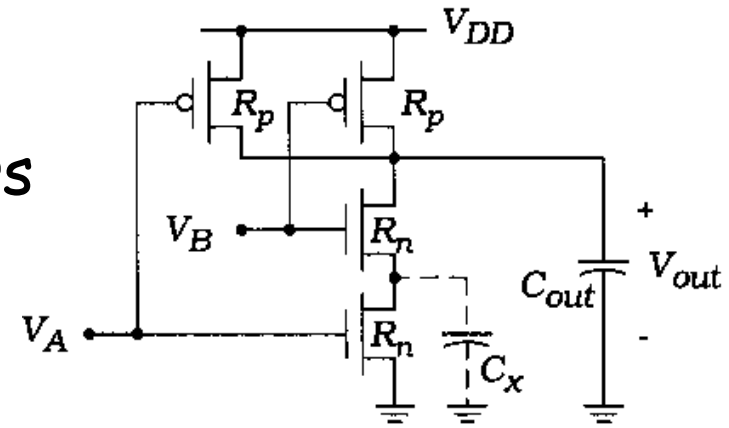



(a) Inverter (b) NAND2 (c) NOR2

# NAND/NOR Transient Summary

- Critical Delay Path
  - paths through series transistors will be slower
  - more series transistors means worse delays

- Tx Sizing Considerations
  - increase W in series transistors
  - balance $\beta_n/\beta_p$ for each cell




- Worst Case Transition
  - when all series transistor go from OFF to ON
  - and all internal caps have to be
    - charged (NOR)
    - discharged (NAND)

# Performance Considerations

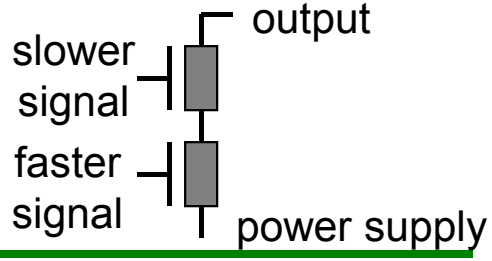
- Speed based on $\beta n$, $\beta p$ and parasitic caps
- DC performance ($V_M$, noise) based on $\beta n/\beta p$
- Design for speed not necessarily provide good DC performance
- Generally set tx size to optimize speed and then test DC characteristics to ensure adequate noise immunity

---

- Review Inverter: Our performance reference point
  - for symmetry ($V_M=V_{DD}/2$), $\beta n = \beta p$
    - which requires $(W/L)_p = \mu_n/\mu_p\ (W/L)_n$
- Use inverter as reference point for more complex gates

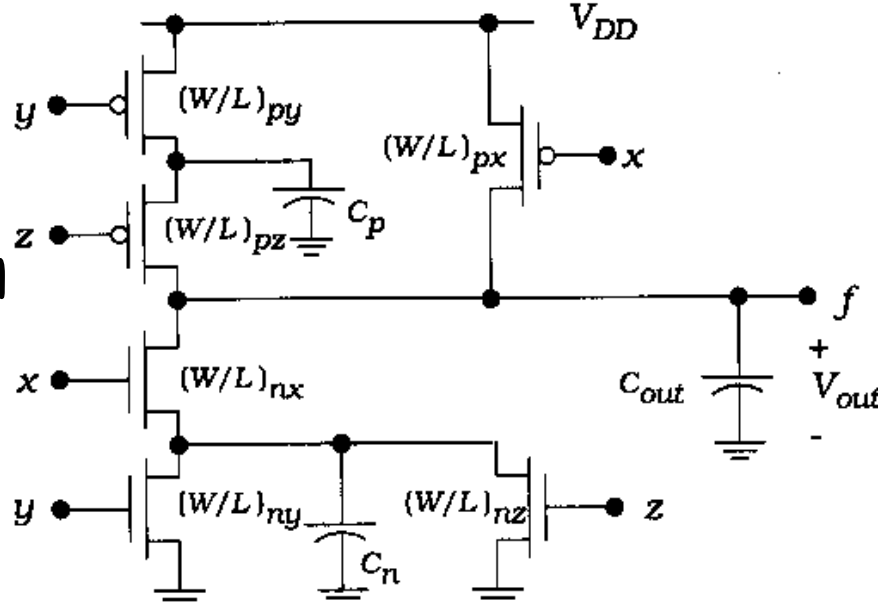
---

- Apply slowest arriving inputs to series node closest to output
  - let faster signals begin to charge/discharge nodes closer to VDD and Ground



# Timing in Complex Logic Gates

- Critical delay path is due to series-connected transistors
- Example: $f = \overline{x\,(y+z)}$
  - assume all tx are same size
- Fall time critical delay
  - worst case, x ON, and y or z ON
  - $t_f = 2.2\,\tau_n$
    - $\tau_n = Rn\,Cn + 2\,Rn\,C_{out}$
      - $C_{out} = 2C_{Dp} + C_{Dn} + C_L$
      - $Cn = 2C_{Dn} + C_{Sn}$
- Rise time critical delay
  - worst case, y and z ON, x OFF
  - $t_r = 2.2\,\tau_p$
    - $\tau_p = Rp\,Cp + 2\,Rp\,C_{out}$
      - $C_{out} = 2C_{Dp} + C_{Dn} + C_L$
      - $Cp = C_{Dp} + C_{Sp}$




size vs. tx speed considerations

⇑Wnx ⇒ ⇓Rn but ⇑Cout and ⇑Cn

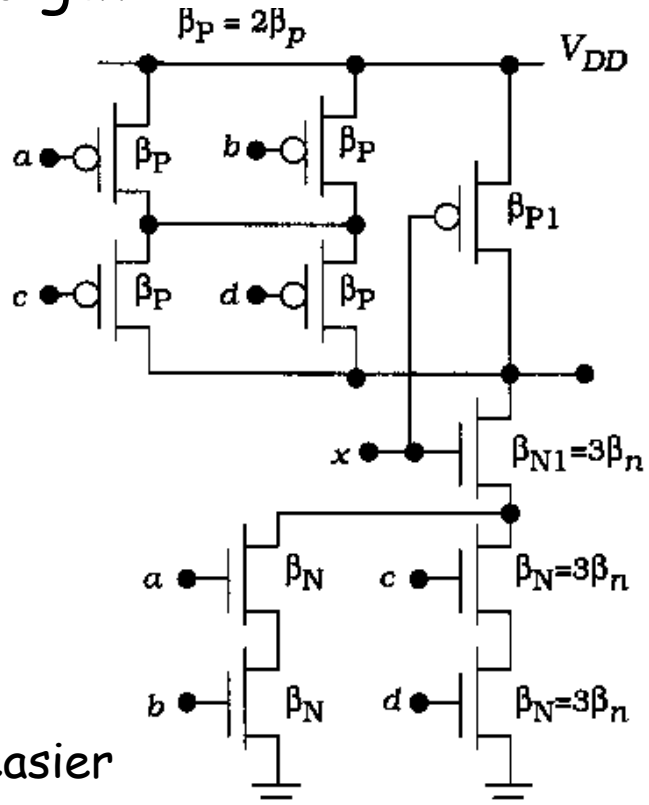
⇓Wny ⇒ ⇓Cn but ⇑Rn

⇑Wpz ⇒ ⇓Rp but ⇑Cout and ⇑Cp

⇓Wpx ⇒ no effect on critical path!

# Sizing in Complex Logic Gates

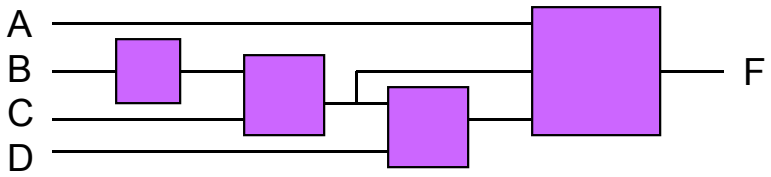
- Improving speed within a single logic gate
- An Example: f=(a b+c d) x
- nMOS
  - discharge through 3 series nMOS
  - set $\beta_N = 3\beta n$
- pMOS
  - charge through 2 series pMOS
  - set $\beta_P = 2\beta p$
  - but, Mxp is alone so $\beta_{P1} = \beta p$
    - but setting $\beta_{P1} = 2\beta p$ might make layout easier




- These large transistors will increase capacitance and layout area and may only give a small increase in speed
- Advanced logic structures are best way to improve speed

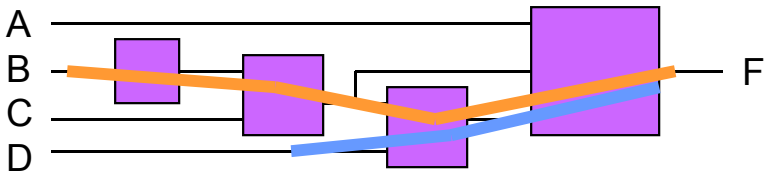
# Timing in Multi-Gate Circuits

- What is the worst-case delay in multi-gate circuits?




| A | B | C | D | F | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | C↑ |
| 0 | 0 | 0 | 1 | 0 | C↑D↓ |
| 0 | 0 | 1 | 0 | 1 | |
| 1 | 0 | 0 | 0 | 0 | B↑ |
| 1 | 1 | 0 | 0 | 1 | |
| 1 | 1 | 1 | 1 | 1 | |

  - too many transitions to test manually
- Critical Path
  - longest delay through a circuit block
  - largest sum of delays, from input to output
  - intuitive analysis: signal that passes through most gates
    - not always true. can be slower path through fewer gates




path through most gates

critical path if delay at D input is very slow

# Power in Multi-Input Logic Gates

- Inverter Power Consumption
  - $P = P_{DC} + P_{dyn} = V_{DD}I_{DDQ} + C_{out}V^2_{DD}f$
    - assumes gates switch output state once per clock cycle, f
- Multi-Input Gates
  - same DC component as inverter, $P_{DC} = V_{DD}I_{DDQ}$
  - for dynamic power, need to estimate "activity" of the gate, how often will the output be switching
  - $P_{dyn} = aC_{out}V^2_{DD}f$, a = activity coefficient
  - estimate activity from truth table
    - $a = p_0p_1$
      - $p_0$ = prob. output is at 0
      - $p_1$ = prob. of transition to 1

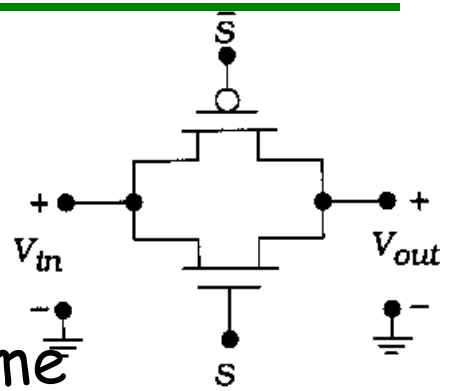
NOR NAND

| A | B | $\overline{A+B}$ | $\overline{A \cdot B}$ |
|---|---|---|---|
| 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 0 | 0 |

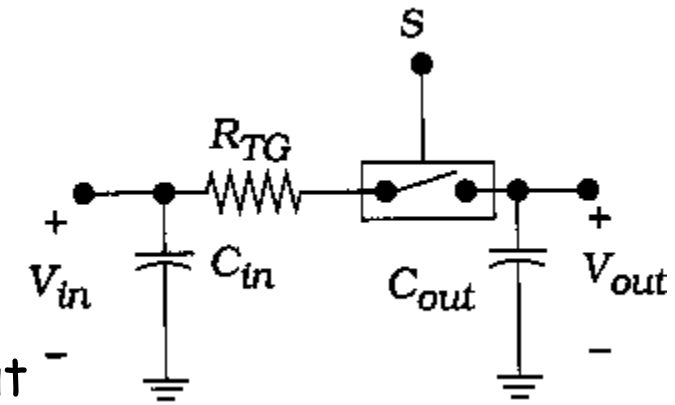
p0=0.75 p0=0.25
p1=0.25 p1=0.75
a=3/16 a=3/16

# Timing Analysis of Transmission Gates

- TG = parallel nMOS and pMOS




- RC Model
  - in general, only one tx active at same time
    - nMOS pulls output low
    - pMOS pushes output high
  - $R_{TG}$ = max (Rn, Rp)
  - Cin = $C_{Sn}$ + $C_{Dp}$
    - if output at higher voltage than input

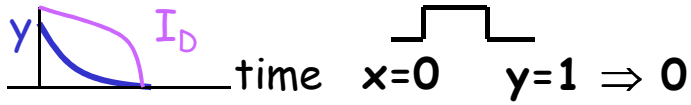


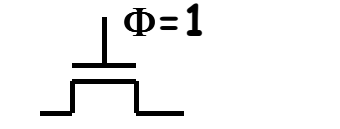

  - larger W will decrease R but increase Cin

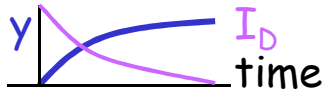
- Note: no connections to VDD-Ground. Input signal, Vin, must drive TG output; TG just adds extra delay

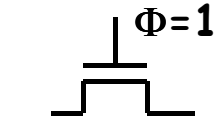
# Pass Transistor

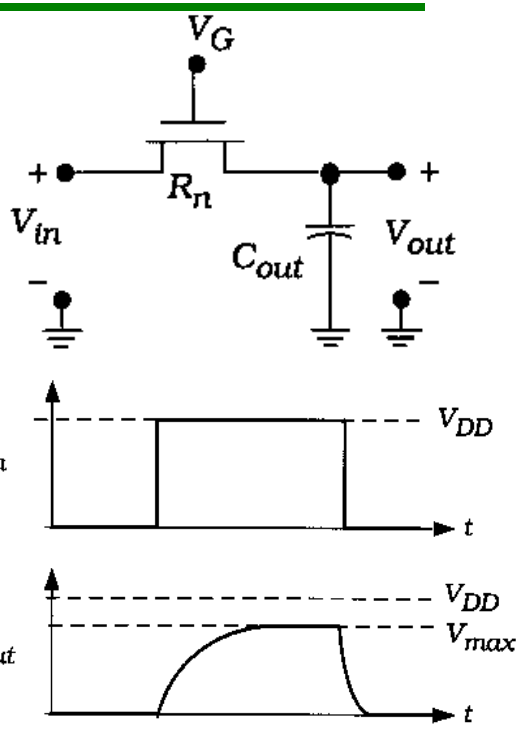
- Single nMOS or pMOS tx
- Often used in place of TGs
  - less area and wiring
  - can't pull to both VDD and Ground
  - typically use nMOS for better speed
- Rise and Fall Times
  - $\tau_n = Rn\, C_{out}$
  - $t_f = 2.94\, \tau_n$
  - $t_r = 18\, \tau_n$
    - much slower than fall time




- nMOS can't pull output to VDD
  - rise time suffers from threshold loss in nMOS