FOURTH EDITION

CMOS VLSI DESIGN

A CIRCUITS AND SYSTEMS PERSPECTIVE

NEIL H. E. WESTE   DAVID MONEY HARRIS

Lecture 5: DC & Transient Response

Outline

- ❑ Pass Transistors
- ❑ DC Response
- ❑ Logic Levels and Noise Margins
- ❑ Transient Response
- ❑ RC Delay Models
- ❑ Delay Estimation

Pass Transistors

- ❑ We have assumed source is grounded
- ❑ What if source > 0?
  - – e.g. pass transistor passing V_{DD}
- ❑ $V_g = V_{DD}$
  - – If $V_s > V_{DD}\text{-}V_t$, $V_{gs} < V_t$
  - – Hence transistor would turn itself off
- ❑ nMOS pass transistors pull no higher than $V_{DD}\text{-}V_{tn}$
  - – Called a degraded "1"
  - – Approach degraded value slowly (low I_{ds})
- ❑ pMOS pass transistors pull no lower than V_{tp}
- ❑ Transmission gates are needed to pass both 0 and 1
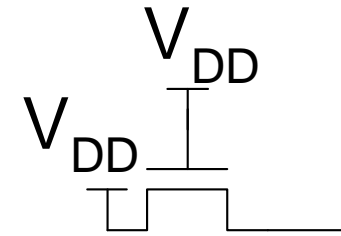
Pass Transistor Ckts

DD

DD

s   DD   tn
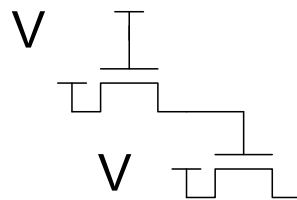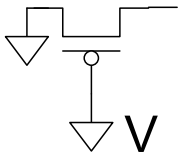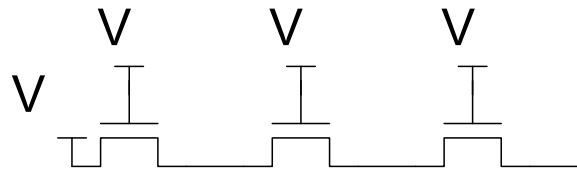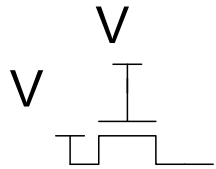
DD

DC Response

- DC Response: V_{out} vs. V_{in} for a gate
- Ex: Inverter
  - When $V_{in} = 0$ -> $V_{out} = V_{DD}$
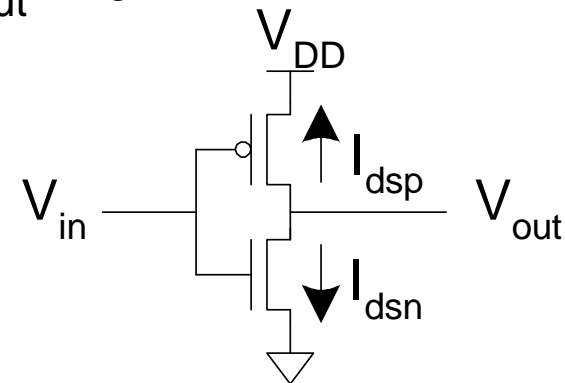 - When $V_{in} = V_{DD}$ -> $V_{out} = 0$
  - In between, V_{out} depends on transistor size and current
  - By KCL, must settle such that $I_{dsn} = |I_{dsp}|$
  - We could solve equations
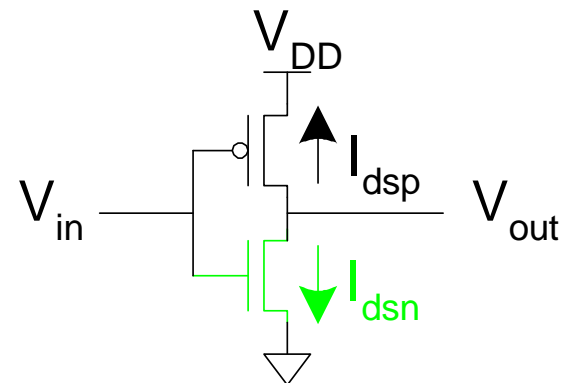 - But graphical solution gives more insight

Transistor Operation

❑ Current depends on region of transistor behavior

❑ For what V_{in} and V_{out} are nMOS and pMOS in
  – Cutoff?
  – Linear?
  – Saturation?

nMOS Operation

Cutoff	Linear	Saturated
$V_{gsn} <$	$V_{gsn} >$	$V_{gsn} >$
	$V_{dsn} <$	$V_{dsn} >$

pMOS Operation

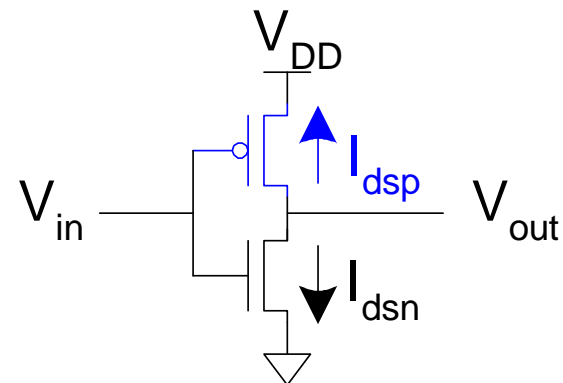| Cutoff	Linear	Saturated
$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
$V_{in} > V_{DD} + V_{tp}$	$V_{in} < V_{DD} + V_{tp}$	$V_{in} < V_{DD} + V_{tp}$
	$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$
	$V_{out} > V_{in} - V_{tp}$	$V_{out} < V_{in} - V_{tp}$

$$V_{gsp} = V_{in} - V_{DD} \qquad V_{tp} < 0$$
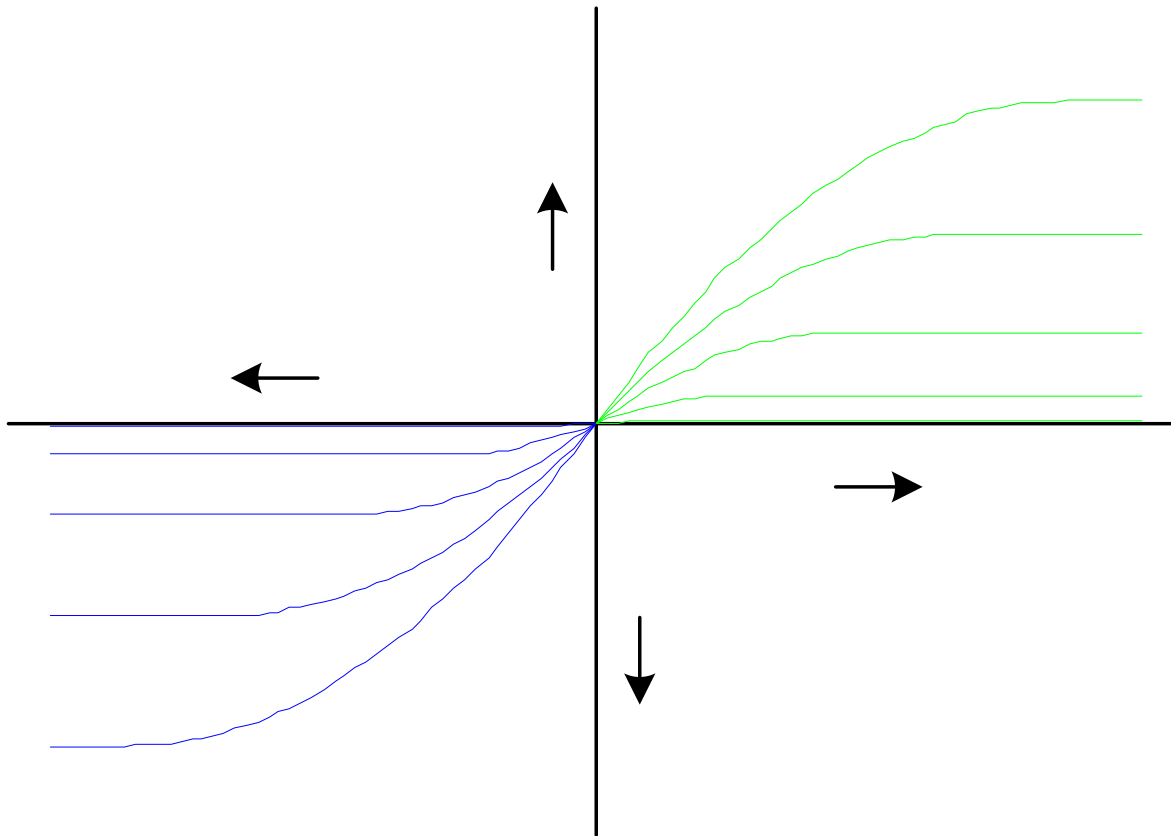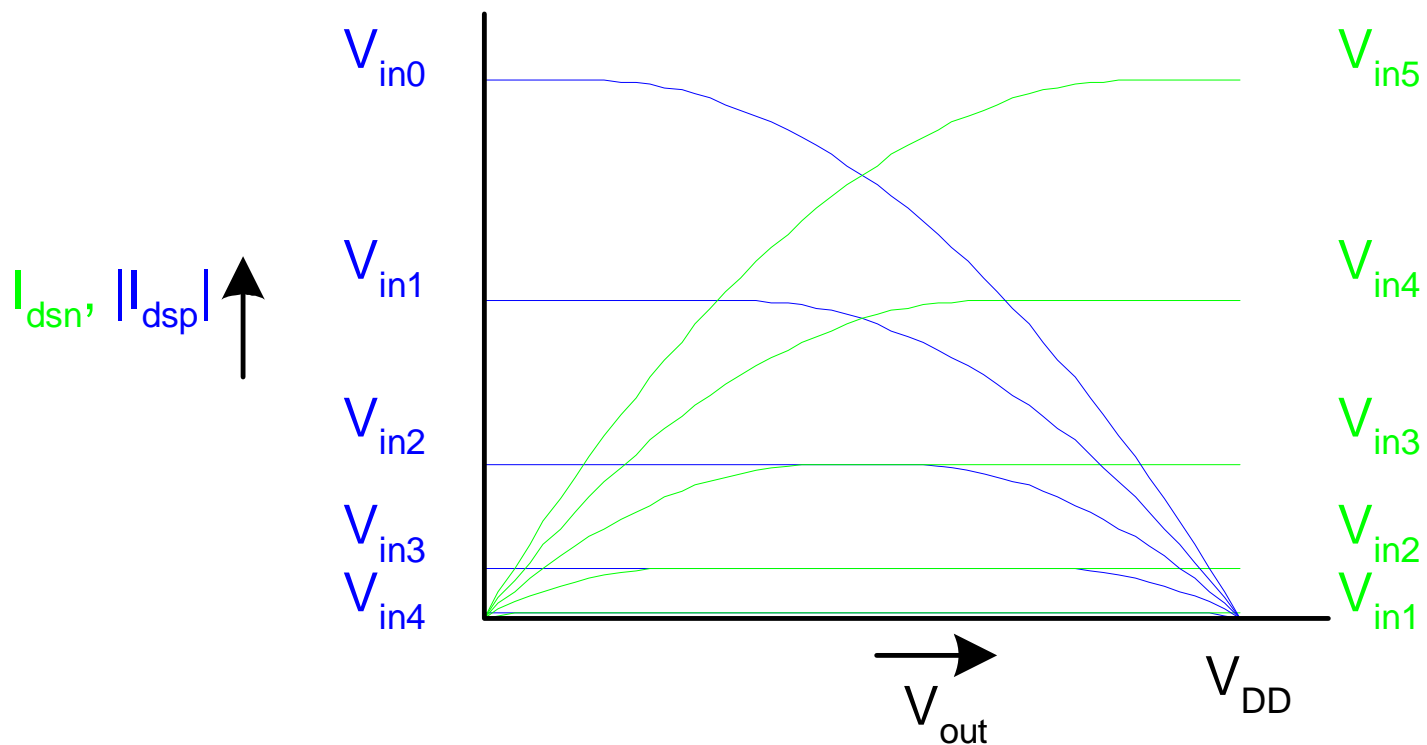
$$V_{dsp} = V_{out} - V_{DD}$$

I-V Characteristics

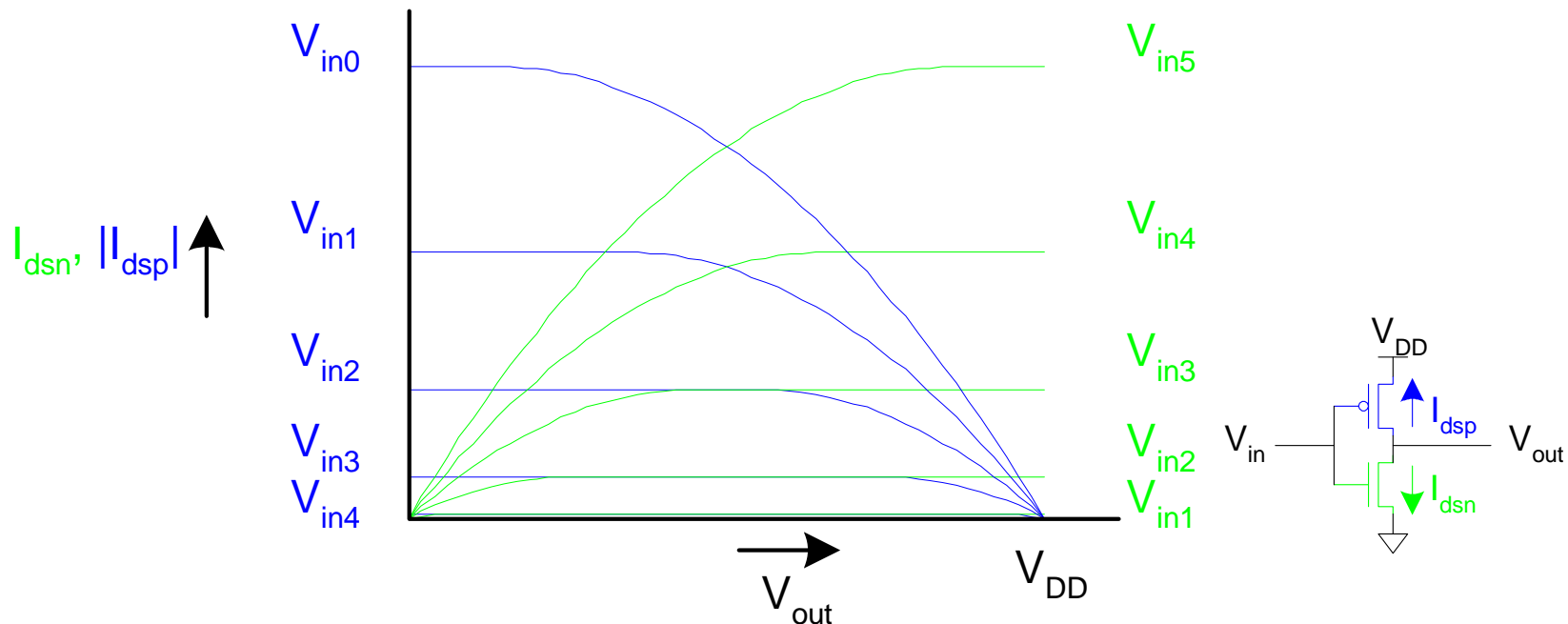❑ Make pMOS is wider than nMOS such that $\beta_n = \beta_p$
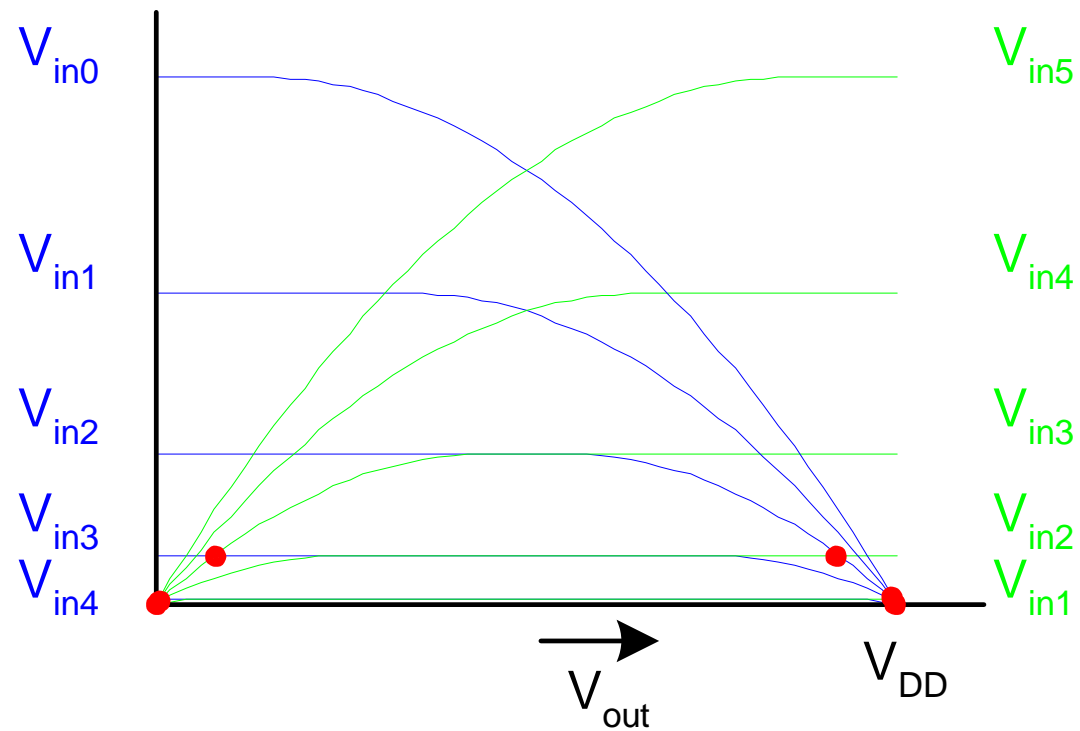
Current vs. V_{out}, V_{in}

Load Line Analysis

□ For a given V_{in}:
  - Plot I_{dsn}, I_{dsp} vs. V_{out}
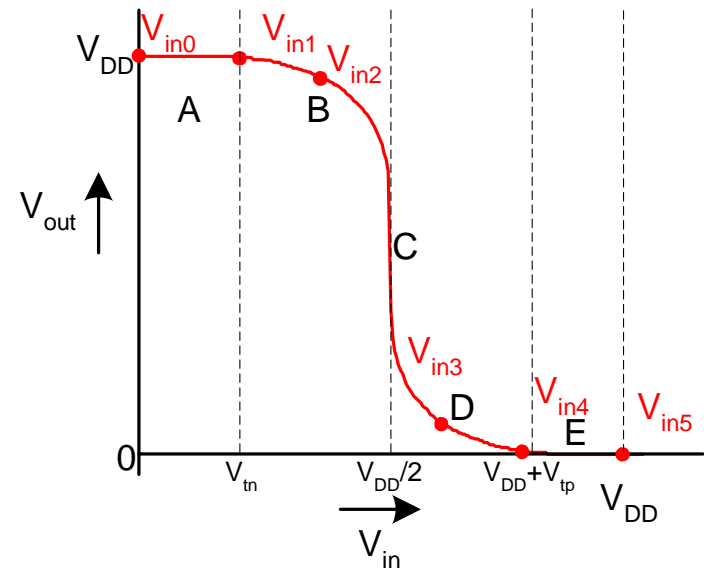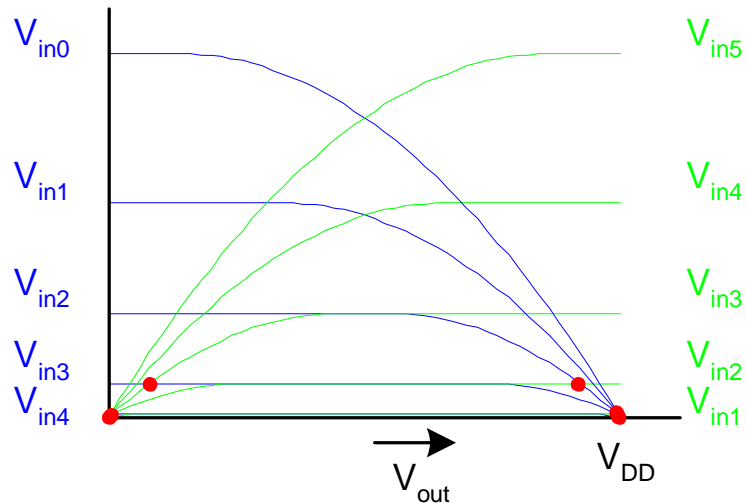 - V_{out} must be where |currents| are equal in

Load Line Analysis
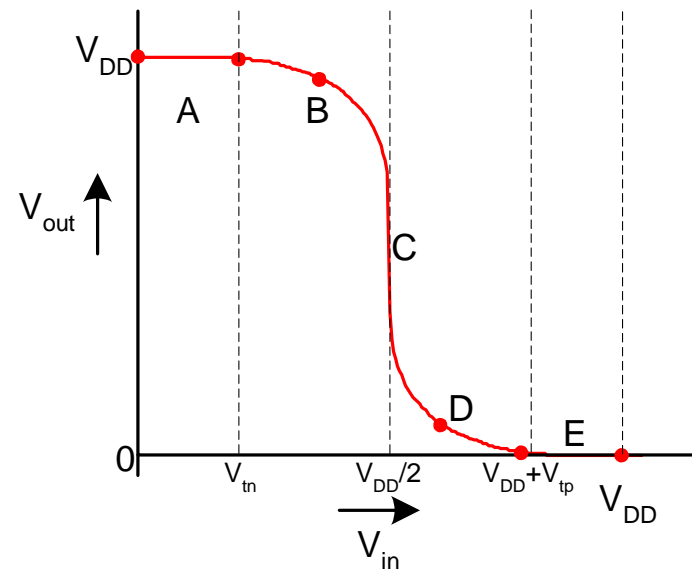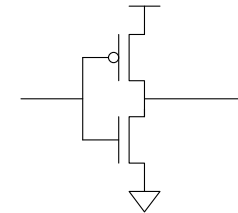
DC Transfer Curve

❑ Transcribe points onto V_{in} vs. V_{out} plot

Operating Regions

❑ Revisit transistor operating regions

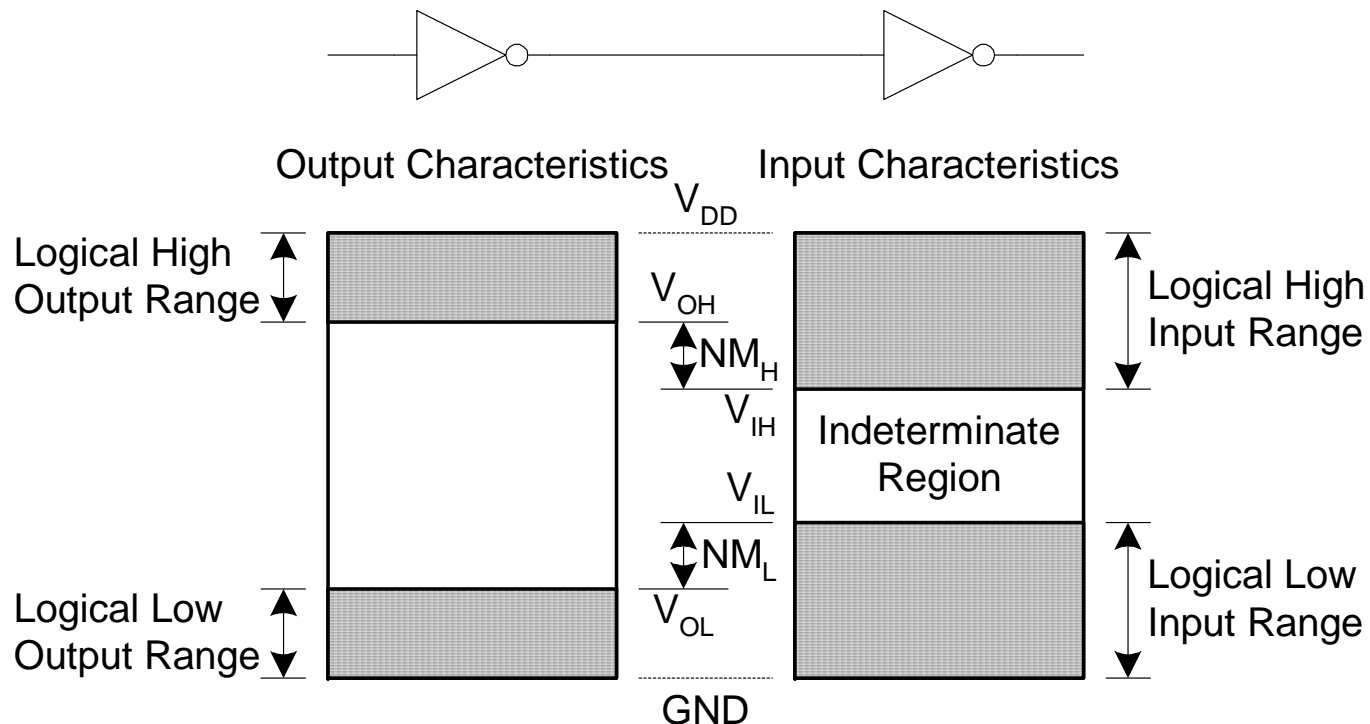| Region	nMOS	pMOS
A		
B		
C		
D		
E		

Beta Ratio

- ❑ If $\beta_p / \beta_n \neq 1$, switching point will move from $V_{DD}/2$
- ❑ Called *skewed* gate
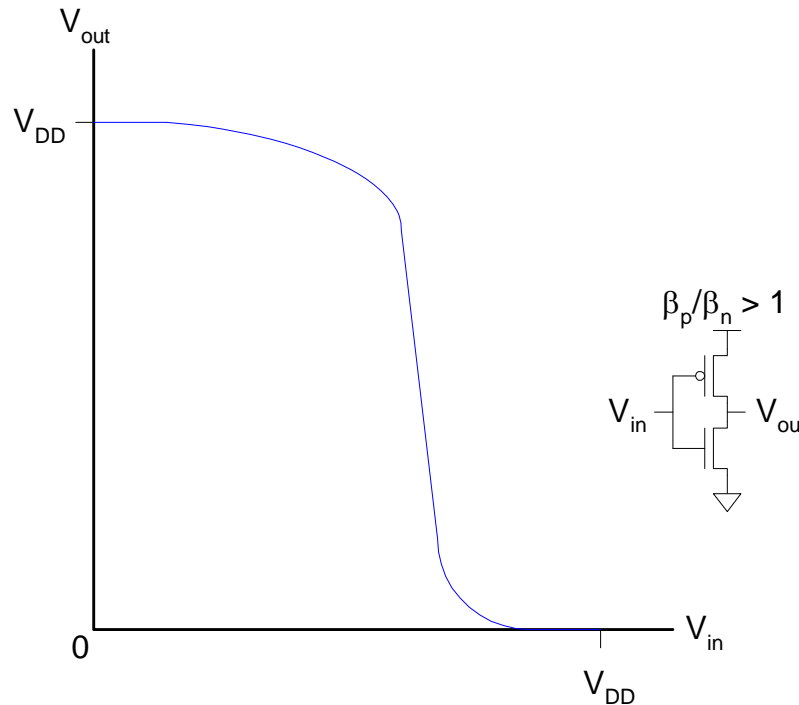- ❑ Other gates: collapse into equivalent inverter
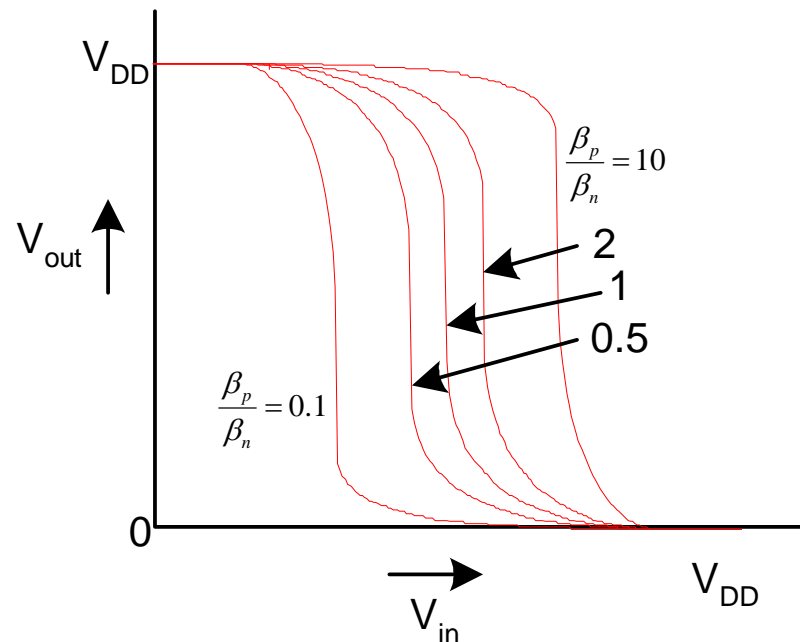
Noise Margins

❑ How much noise can a gate input see before it does not recognize the input?

Logic Levels

❑ To maximize noise margins, select logic levels at
  – unity gain point of DC transfer characteristic

Transient Response

❑ *DC analysis* tells us V_{out} if V_{in} is constant

❑ *Transient analysis* tells us $V_{out}(t)$ if $V_{in}(t)$ changes

– Requires solving differential equations

❑ Input is usually considered to be a step or ramp
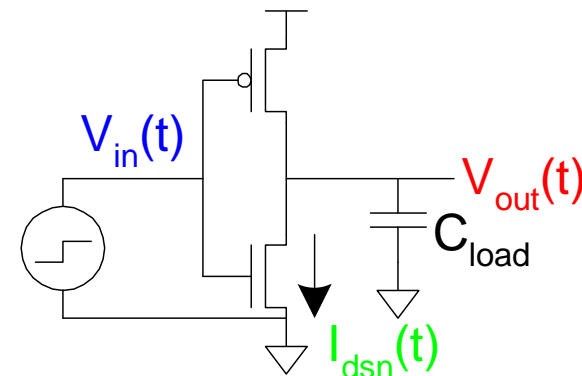
– From 0 to V_{DD} or vice versa

Inverter Step Response

❑ Ex: find step response of inverter driving load cap

$$V_{in}(t) =$$

$$V_{out}(t < t_0) =$$

$$\frac{dV_{out}(t)}{dt} =$$

$$I_{dsn}(t) = \begin{cases} & t \le t_0 \\ & V_{out} > V_{DD} - V_t \\ & V_{out} < V_{DD} - V_t \end{cases}$$

Delay Definitions

- **t_{pdr}**: *rising propagation delay*
  - From input to rising output crossing $V_{DD}/2$
- **t_{pdf}**: *falling propagation delay*
  - From input to falling output crossing $V_{DD}/2$
- **t_{pd}**: *average propagation delay*
  - $t_{pd} = (t_{pdr} + t_{pdf})/2$
- **t_r**: *rise time*
  - From output crossing 0.2 V_{DD} to 0.8 V_{DD}
- **t_f**: *fall time*
 - From output crossing 0.8 V_{DD} to 0.2 V_{DD}
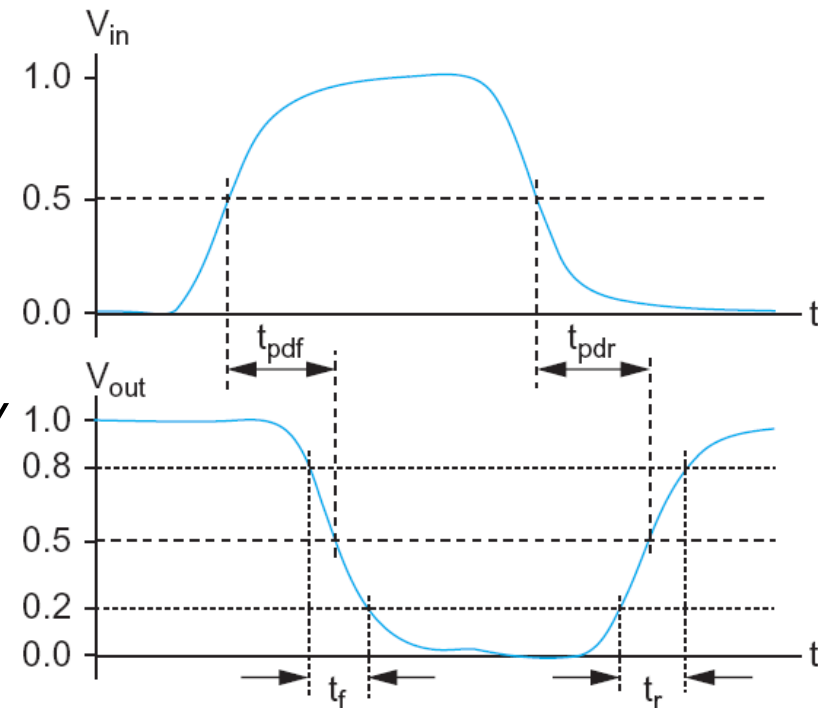
Delay Definitions

- **t_{cdr}**: *rising contamination delay*
  - From input to rising output crossing $V_{DD}/2$
- **t_{cdf}**: *falling contamination delay*
  - From input to falling output crossing $V_{DD}/2$
- **t_{cd}**: *average contamination delay*
  - $t_{pd} = (t_{cdr} + t_{cdf})/2$

Simulated Inverter Delay

- ❑ Solving differential equations by hand is too hard
- ❑ SPICE simulator solves the equations numerically
  - – Uses more accurate I-V models too!
- ❑ But simulations take time to write, may hide insight
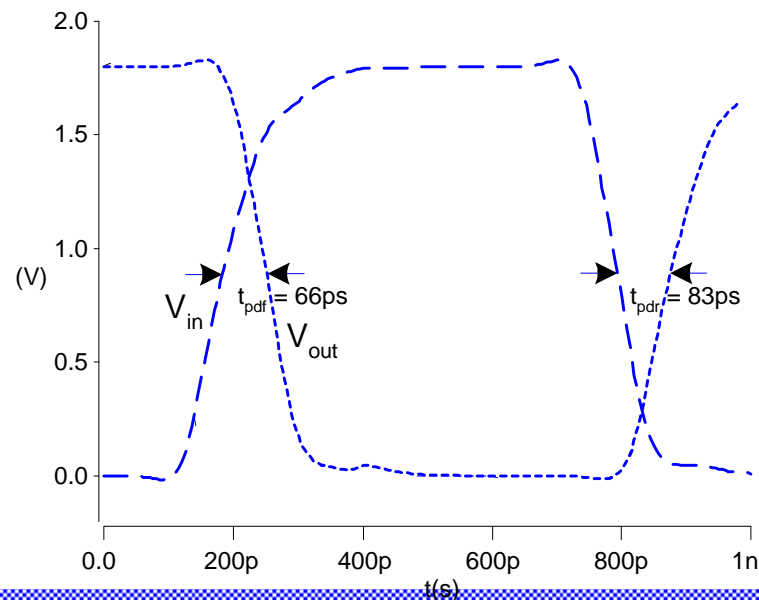
Delay Estimation

- ❑ We would like to be able to easily estimate delay
  - – Not as accurate as simulation
  - – But easier to ask "What if?"
- ❑ The step response usually looks like a 1st order RC response with a decaying exponential.
- ❑ Use RC delay models to estimate delay
  - – C = total capacitance on output node
  - – Use *effective resistance* R
  - – So that $t_{pd} = RC$
- ❑ Characterize transistors by finding their effective R
  - – Depends on average current as gate switches

Effective Resistance

❑ Shockley models have limited value
  – Not accurate enough for modern transistors
  – Too complicated for much hand analysis

❑ Simplification: treat transistor as resistor
  – Replace $I_{ds}(V_{ds}, V_{gs})$ with effective resistance R
    • $I_{ds} = V_{ds}/R$
  – R averaged across switching of digital gate

❑ Too inaccurate to predict current at any given time
  – But good enough to predict RC delay

RC Delay Model

❑ Use equivalent circuits for MOS transistors
  – Ideal switch + capacitance and ON resistance
  – Unit nMOS has resistance R, capacitance C
  – Unit pMOS has resistance 2R, capacitance C
❑ Capacitance proportional to width
❑ Resistance inversely proportional to width
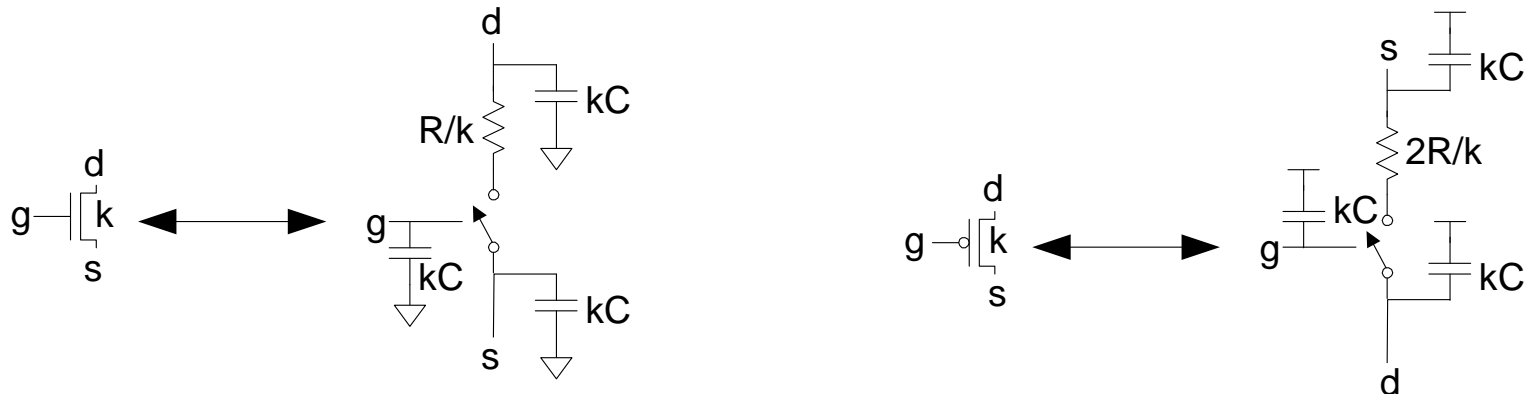
RC Values

- ❑ Capacitance
  - – $C = C_g = C_s = C_d = 2$ fF/μm of gate width in 0.6 μm
  - – Gradually decline to 1 fF/μm in nanometer techs.
- ❑ Resistance
  - – $R \approx 6$ KΩ*μm in 0.6 μm process
  - – Improves with shorter channel lengths
- ❑ Unit transistors
  - – May refer to minimum contacted device (4/2 λ)
  - – Or maybe 1 μm wide device
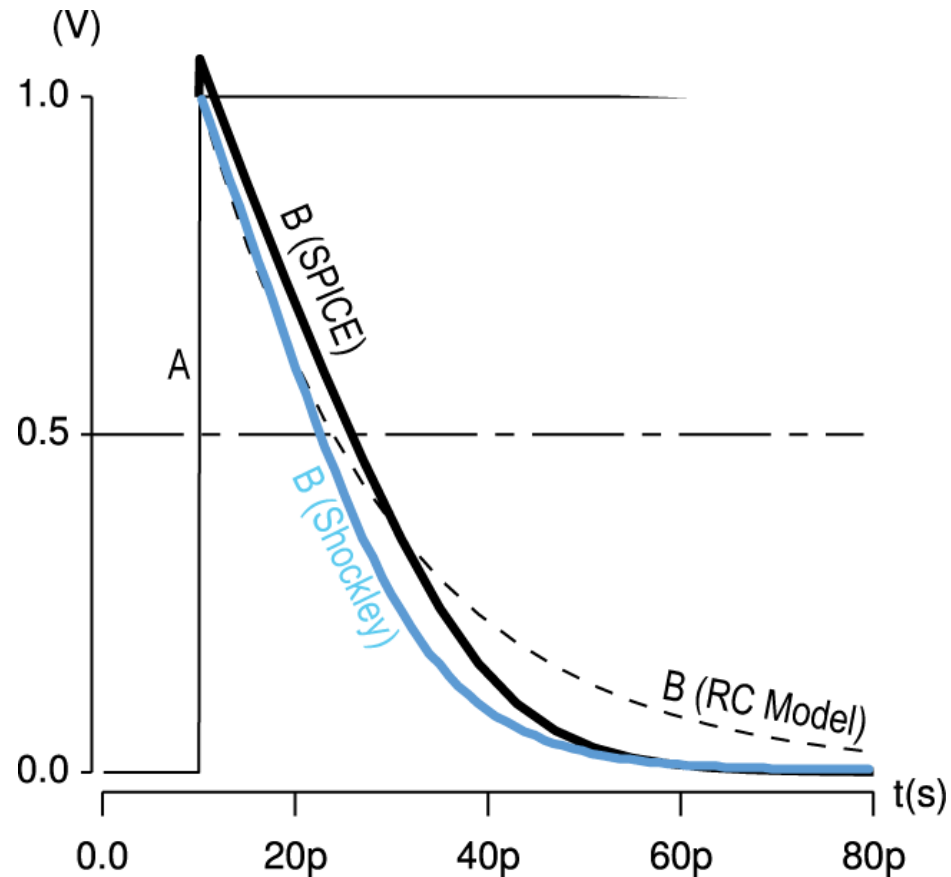 - – Doesn't matter as long as you are consistent

Inverter Delay Estimate
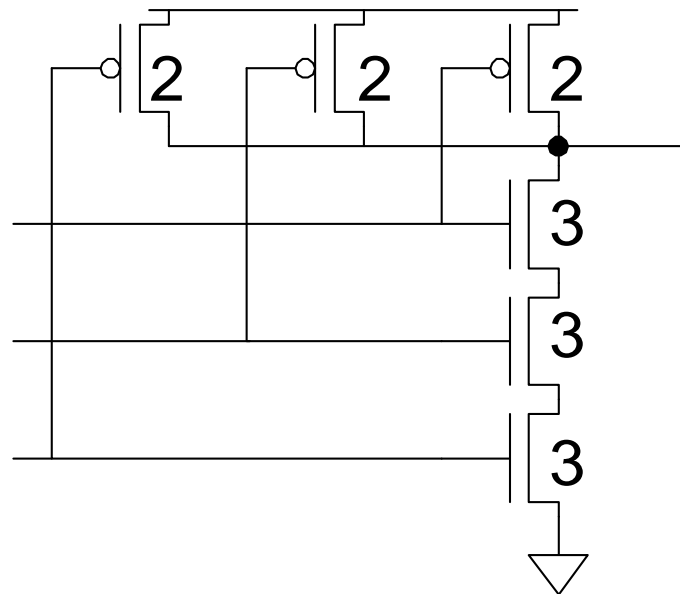
❑ Estimate the delay of a fanout-of-1 inverter
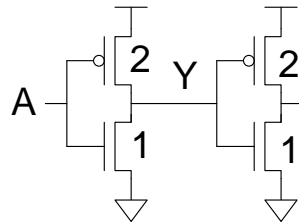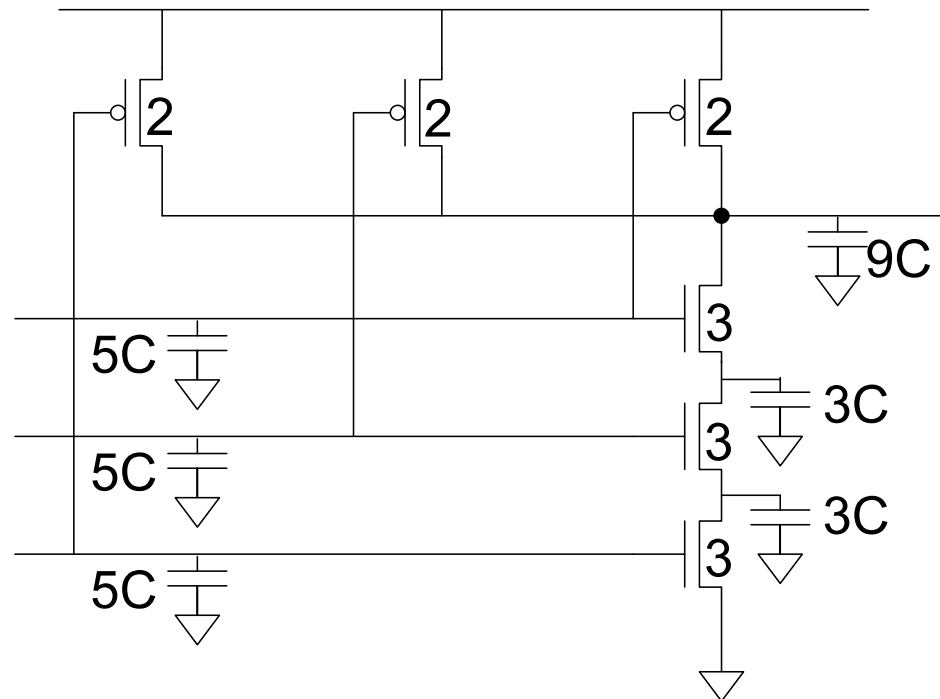

d = 6RC

Delay Model Comparison

Example: 3-input NAND

❑ Sketch a 3-input NAND with transistor widths chosen to achieve effective rise and fall resistances equal to a unit inverter (R).

3-input NAND Caps

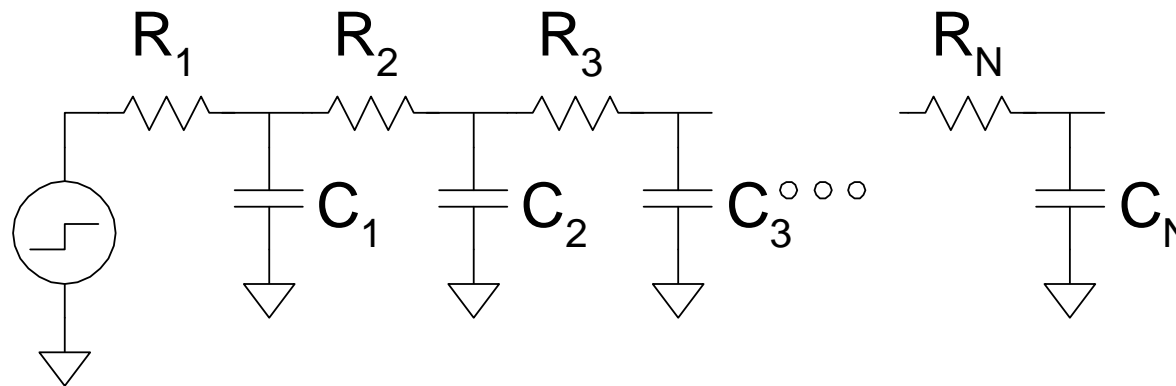❑ Annotate the 3-input NAND gate with gate and diffusion capacitance.

Elmore Delay

- ❑ ON transistors look like resistors
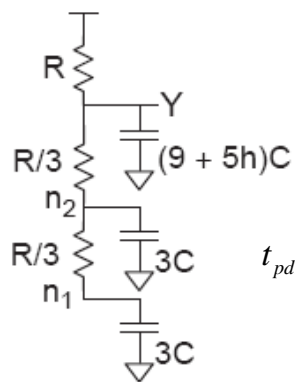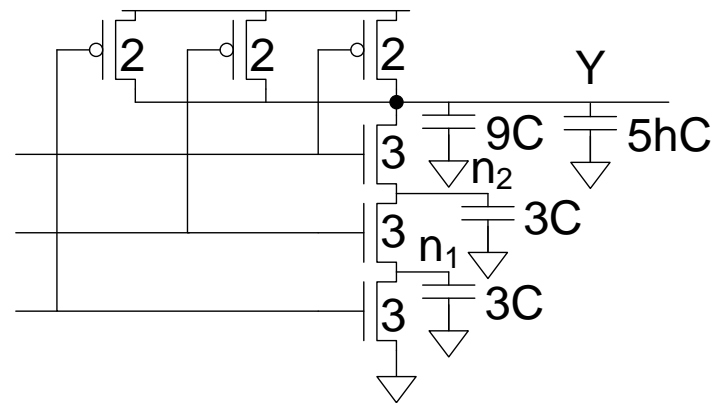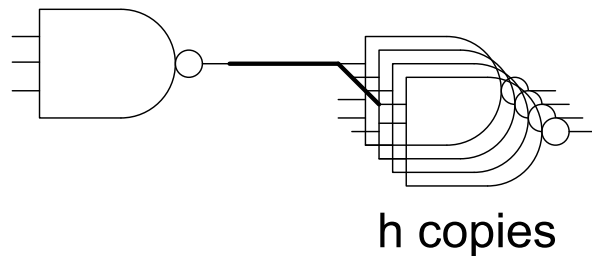- ❑ Pullup or pulldown network modeled as *RC ladder*
- ❑ Elmore delay of RC ladder

$$t_{pd} \approx \sum_{\text{nodes } i} R_{i-to-source} C_i$$

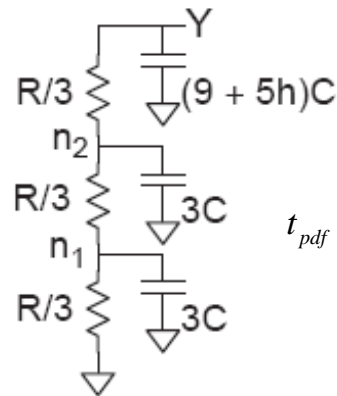$$= R_1 C_1 + (R_1 + R_2) C_2 + ... + (R_1 + R_2 + ... + R_N) C_N$$

Example: 3-input NAND

❑ Estimate worst-case rising and falling delay of 3-input NAND driving h identical gates.

h copies

$$t_{pdr} = (9 + 5h)RC$$

$$t_{pdf} = (3C)\left(\tfrac{R}{3}\right) + (3C)\left(\tfrac{R}{3} + \tfrac{R}{3}\right) + \left[(9+5h)C\right]\left(\tfrac{R}{3} + \tfrac{R}{3} + \tfrac{R}{3}\right)$$
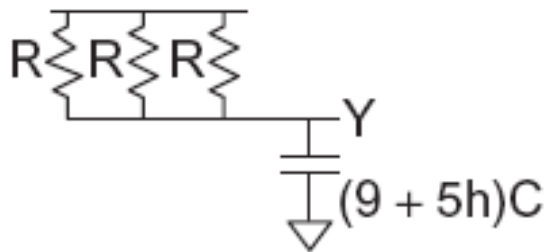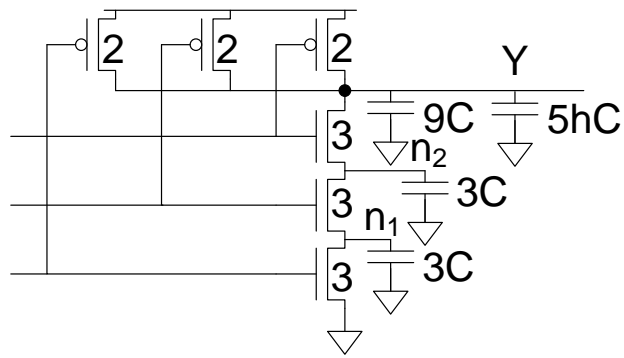$$= (11 + 5h)RC$$

Delay Components

❑ Delay has two parts

– *Parasitic delay*

• 9 or 11 RC

• Independent of load

– *Effort delay*

• 5h RC

• Proportional to load capacitance

Contamination Delay

❑ Best-case (contamination) delay can be substantially less than propagation delay.

❑ Ex: If all three inputs fall simultaneously

$$t_{cdr} = \left[\left(9 + 5h\right)C\right]\left(\frac{R}{3}\right) = \left(3 + \frac{5}{3}h\right)RC$$

Diffusion Capacitance

- ❑ We assumed contacted diffusion on every s / d.
- ❑ Good layout minimizes diffusion area
- ❑ Ex: NAND3 layout shares one diffusion contact
  - – Reduces output capacitance by 2C
  - – Merged uncontacted diffusion might help too
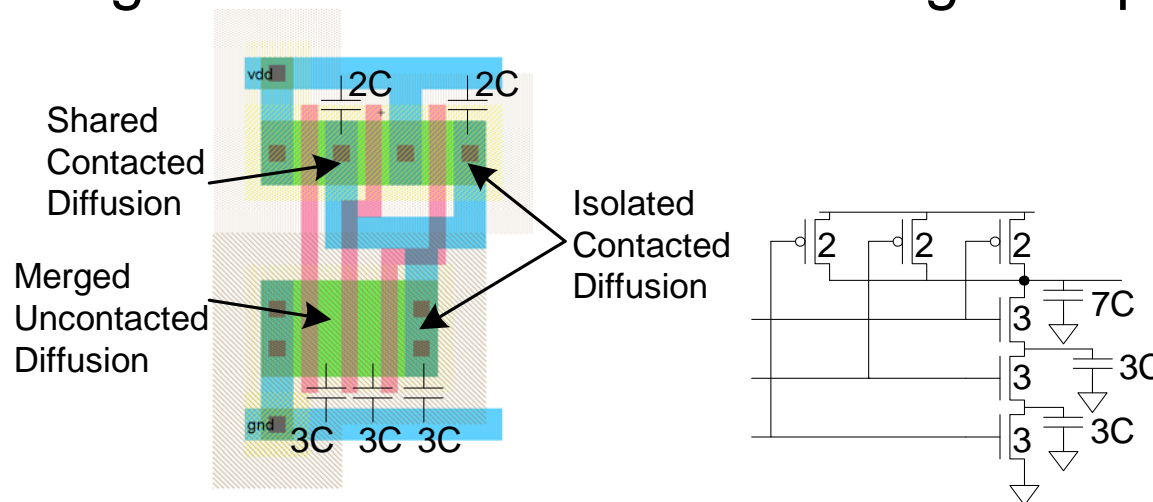
Layout Comparison

❑ Which layout is better?