Lecture 24

CMOS Capacitance and Circuit Delay

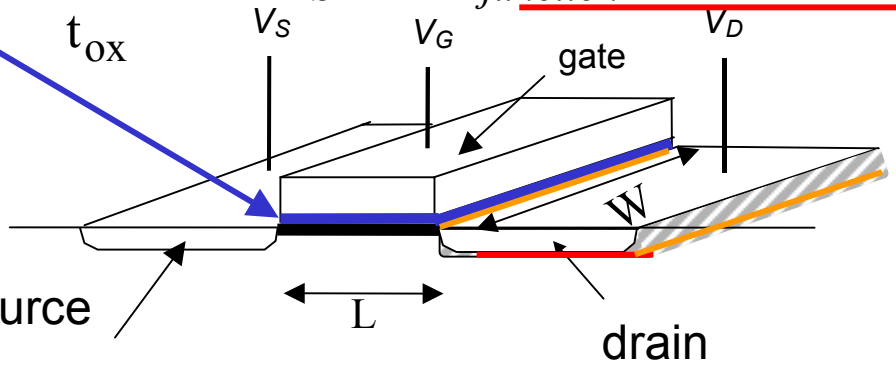
A) CMOS Structure and Capacitance
B) Gate and Source Drain Capacitance Model
C) Cascade Inverter Delay
D) Capacitance from Logic Function
E) Fan-Out and Logic Delay

Reading: Schwarz and Oldham, pp. 518-526, and lectures 16-19.

Dynamic Performance: Add Capacitances

$$C_G = C_{ox} \times W \times L = \frac{\varepsilon_{ox} WL}{t_{ox}}$$

$$C_{SD} = C_{junction} \times W \times 4L_{min} + C_{perimeter} \times W$$



Capacitance C_G is between gate and the underlying channel, which is connected to the source, $C_{GS} = C_G$ and hence is modeled as capacitance to a.c. ground.

Capacitance C_{SD} has a bottom and out-side perimeter between the source or drain and the underlying substrate which is connected to a.c. ground. There is also a gate perimeter component for which there is a 2X magnifying (Miller) effect on the S/D side because the gate swings the opposite direction of the S/D.

Gate Capacitance Model

$$C_G = C_{ox} \times W \times L = \frac{\varepsilon_{ox} WL}{t_{ox}}$$

Dielectric Constant for SiO_2

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \frac{(8.85x10^{-14}\,F/cm)(3.9)}{6x10^{-7}\,cm} = 5.75x10^{-7}\,F/cm^2$$

L = 0.25μm **W = 0.375μm**

$$C_G = (5.75x10^{-7}\,F/cm^2) \times 0.375x10^{-4} \times 0.25x10^{-4} = 0.54\,fF$$

EE42 Simplified C_G model:

$C_{G/MS}$ = 0.4fF/MS (ie. 0.4 fF per Minimum Square of gate layout)

Example For Gates with the minimum length of 0.25 μm:

$(W/L)_n$ = 0.375/0.25 = 1.5 has 1.5 MS and has C_G = 0.6 fF.

$(W/L)_p$ = 0.75/0.25 = 3.0 has 3.0 MS and has C_G = 1.2 fF.

Note: C_G is proportional to (W/L)

Source/Drain Capacitance Model

$$C_{SD} = C_{junction} \times W \times 4L_{\min} + C_{perimeter} \times W$$

Accurately evaluating C_{SD} involves modeling the advanced physics of junctions and is considered in somewhat more detail in EECS 141.

C_{SD} is generally proportional to W.

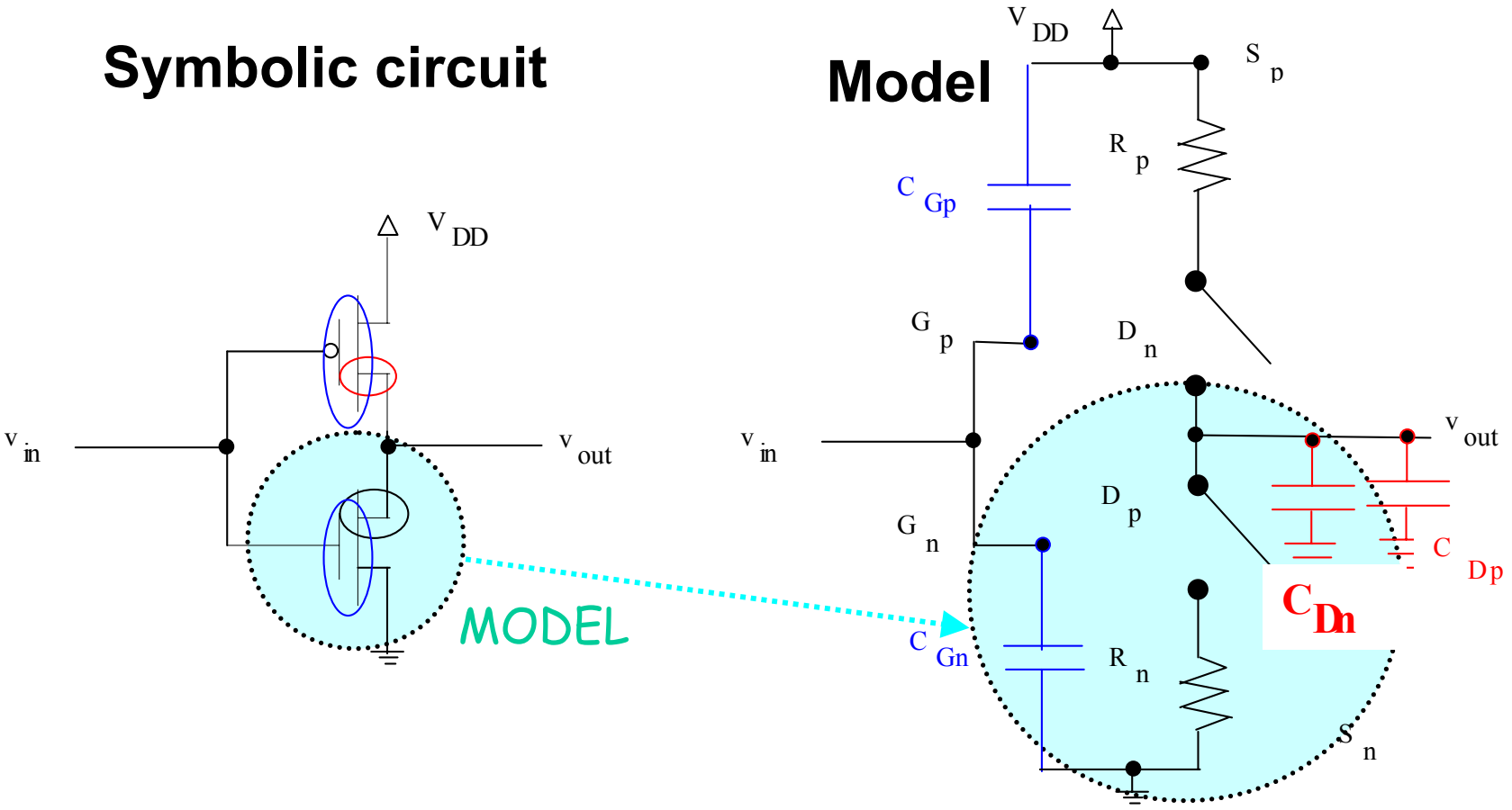
EE42 Simplified C_G model:

C_{SD} = 0.4fF/L_{MIN} (ie. 0.4 fF per Minimum 0.25μm Feature length)

Example: $(W/L)_n$ = 0.375/0.25, => 1.5L with C_{SD} = 0.6 fF.

Example: $(W/L)_p$ = 0.75/0.25 => 3.0L with C_G = 1.2 fF.

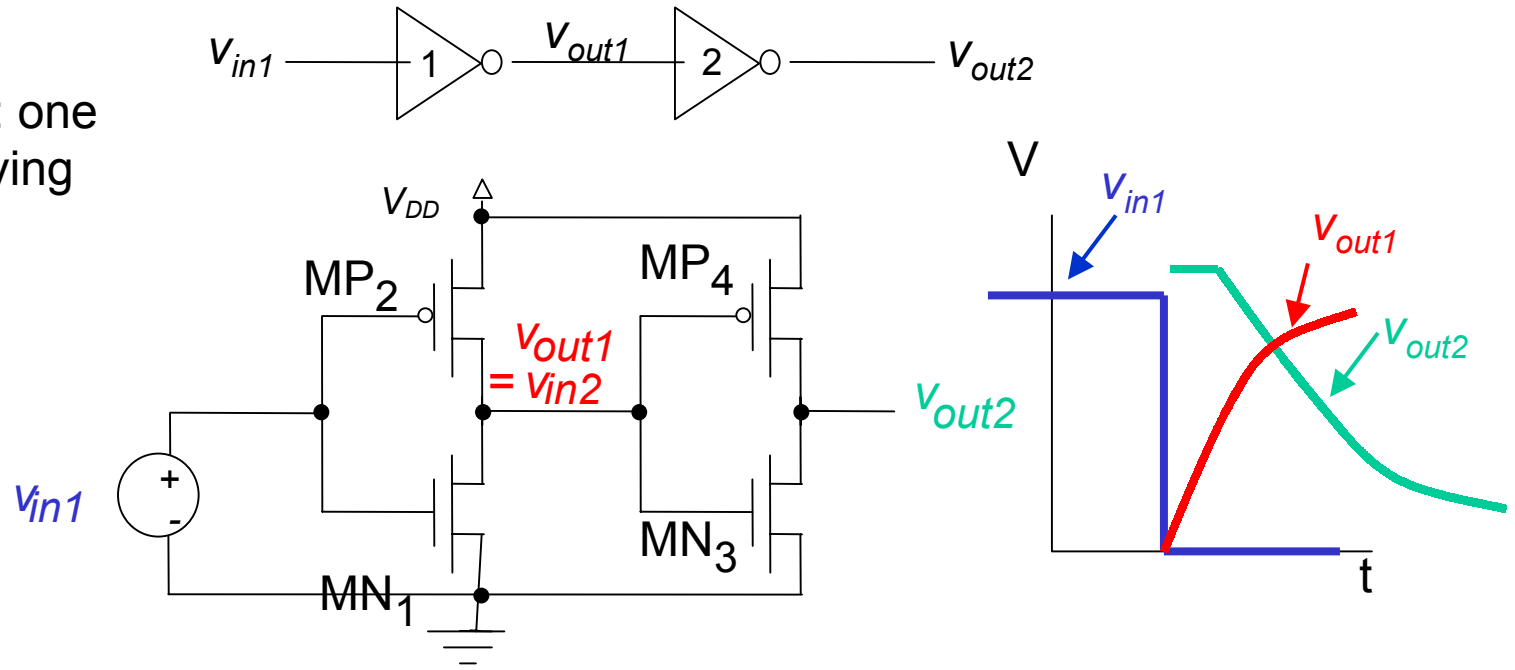
The CMOS Inverter with Parasitic capacitances



Note that the switches are NOT independent , in fact they are “ganged”

Gate-Delay Analysis -- Identify key Components

Basic case: one inverter driving another



Suppose V_{in1} goes from high to low. → MP_2 turns on and MN_1 turns off.

Then V_{out1} goes from low to high (but a little bit later … i.e. delayed).

Of course V_{in2} is the same as V_{out1} .

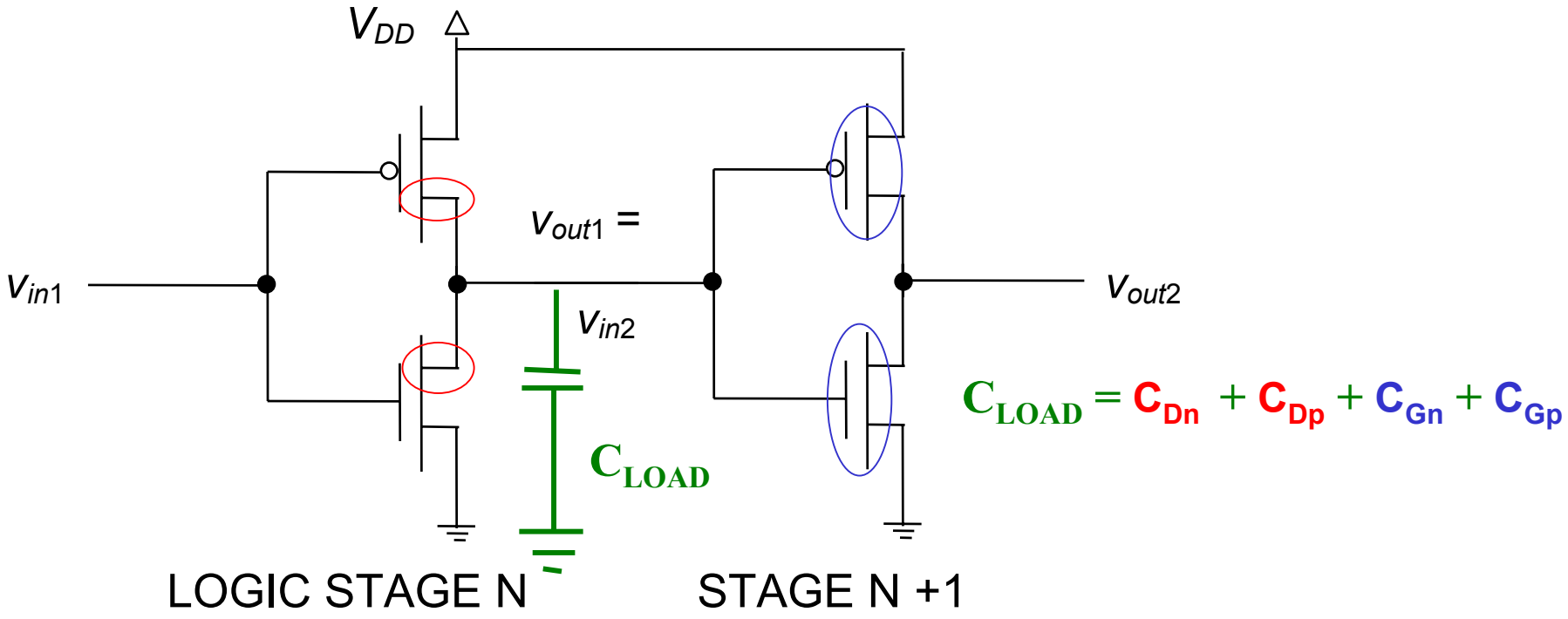
Thus V_{out2} goes from high to low (delayed even more from the input V_{in1}).

"Cascaded" CMOS Inverters

Note that there are no resistors, capacitors, inductors in a CMOS circuit -- there are only NMOS and PMOS transistors.

The output of stage 1 must charge the Source/Drain capacitances of the first stage and the gate capacitances of the second stage.

That is **C_{Dn} and C_{Dp} of the first stage and C_{Gn} and C_{Gp} of the second stage.**



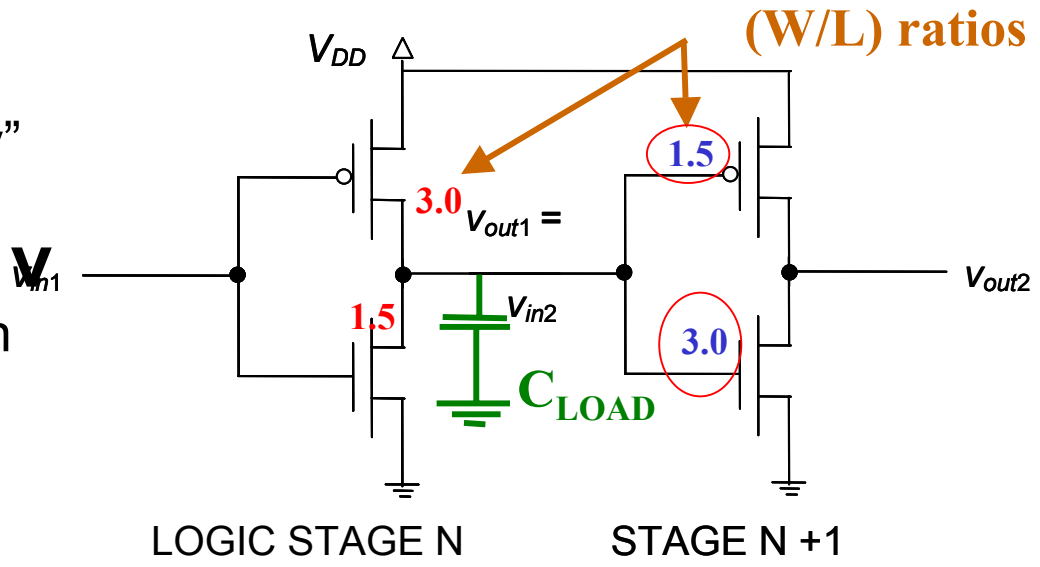
Typical Capacitance Values:

Consider "0.25µm technology"

with a typical NMOS device

0.25 X 0.375 µm as pull-down

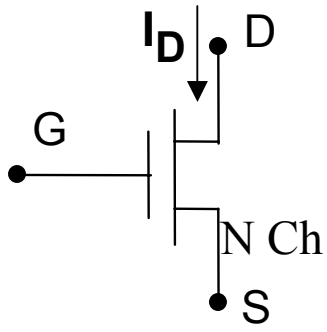
and 0.25 X 0.75 pull-up



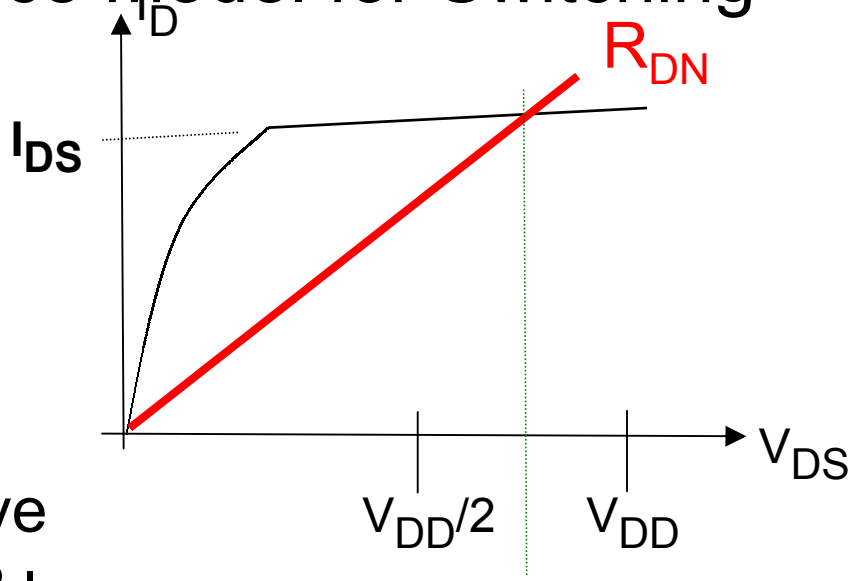
$$C_{LOAD} = C_{Dn} + C_{Dp} + C_{Gn} + C_{Gp}$$

$$C_{LOAD} = (1.5 + 3.0 + 1.5 + 3.0)(0.4fF) = (9)(0.4fF) = 3.6fF$$

Resistance Model for Switching



The circuit symbol



From $I_{DS} = CdV/dt$ we have

$\Delta t = C\Delta V / I_{DS} = C V_{DD} / 2 I_{DS}$

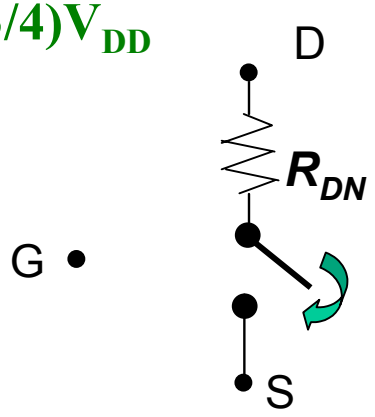
But if we had an RC discharge

$\Delta t = 0.69RC$ so the effective resistance, R_{DN} ,$= \Delta t / 0.69C$

$= V_{DD}/(0.69 \text{ X } 2 I_{DS}) = .72 V_{DD} / I_{DS}$

So we use $$R_{DN} \approx \left(\frac{3}{4}\right)\frac{V_{DD}}{I_{DS}}$$

$(3/4)V_{DD}$



Electrical Model

Resistance Model for Switching

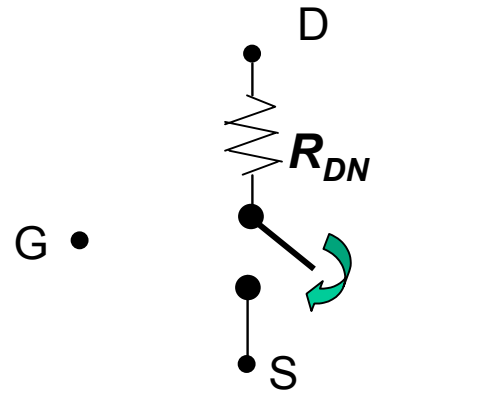
For the 0.25 μm technology with V_{DD} =2.5V

For W_p = 0. 75μm

$$R_{DP} \approx \left(\frac{3}{4}\right)\frac{V_{DD}}{I_{DS}} = \left(\frac{3}{4}\right)\frac{2.5V}{158\mu A} = 11.87k\Omega \approx 12k\Omega$$

For W_n = 0.375μm

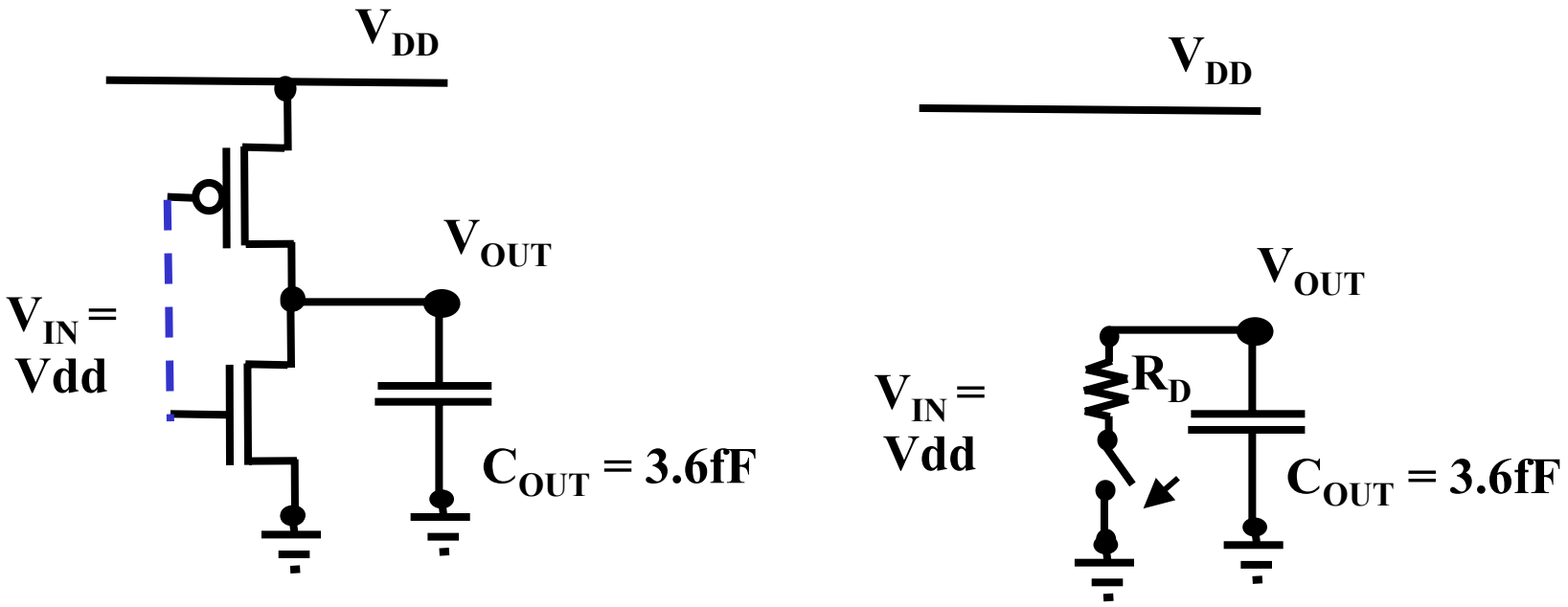
$$R_{DN} \approx \left(\frac{3}{4}\right)\frac{V_{DD}}{I_{DS}} = \left(\frac{3}{4}\right)\frac{2.5V}{196\mu A} = 9.57k\Omega \approx 10k\Omega$$



Electrical Model

Inverter Propagation Delay

Discharge (pull-down)



$$\Delta t = 0.69 R_{DN} C_{LOAD} = 0.69(10k\Omega)(3.6fF) = 25\ ps$$

Discharge (pull-up)

$$\Delta t = 0.69 R_{DP} C_{LOAD} = 0.69(12k\Omega)(3.6fF) = 30\ ps$$

Computer Circuits Are and Order of Magnitude Slower Than a Simple Inverter.

The Source/drain capacitance of a circuit increases roughly proportional to the number of input signals.

The resistance from the output node to ground of V_{DD} increases roughly proportional to the number of input signals.

Slide 13 will give an example

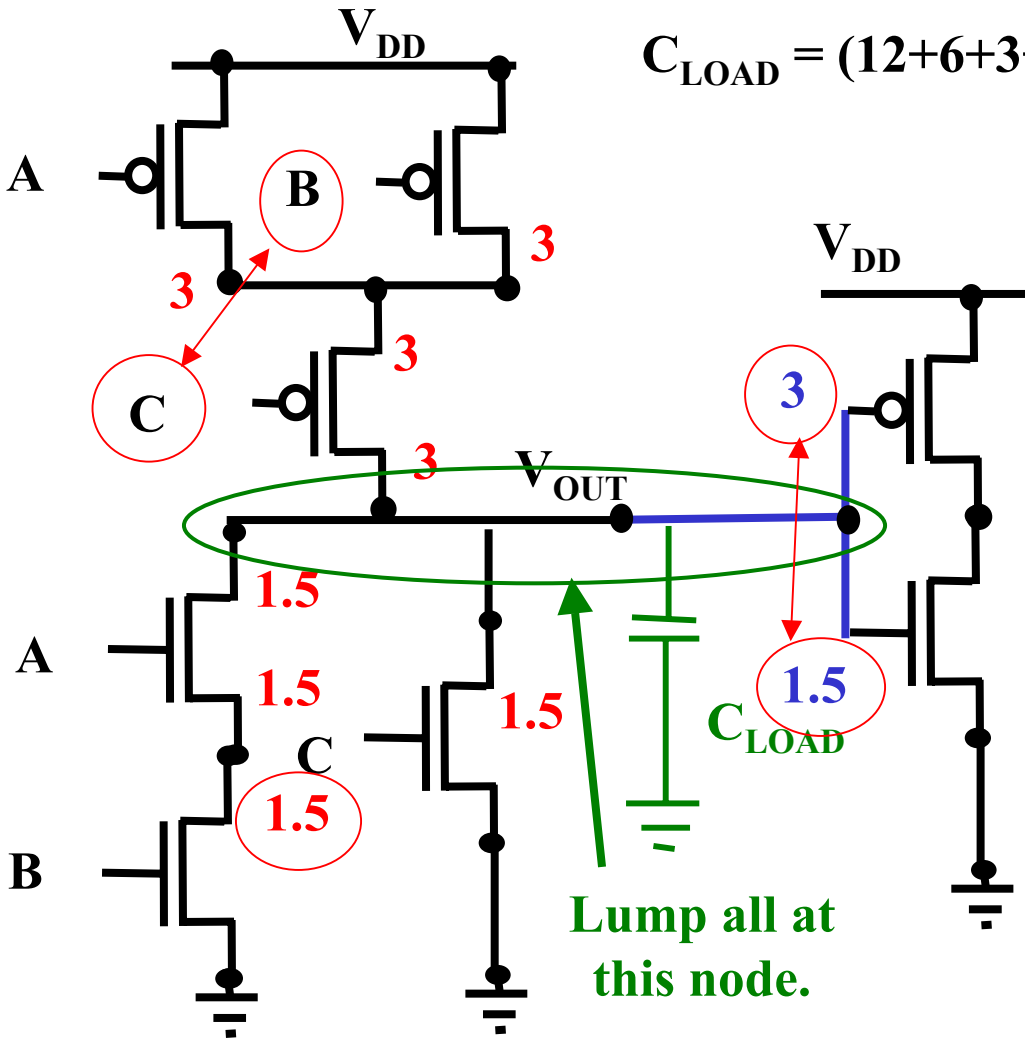
The time to charge the internal source/drain capacitances is the product of the capacitance and resistance and roughly increases as the square of the number of input signals.

The delivery of the output signal to more than one logic gate (Fan-Out) introduces further loading proportional to the Fan-Out.

Study this page carefully as three starting point mistakes were corrected.

Version Date 12/01/01

Example CMOS Circuit



$C_{LOAD} = (12+6+3+1.5)(0.4fF) = 9$ fF

Worst case is a=1, c =0, and b changes 1 => 0

$R = 2R_{DP} = 24k\Omega$

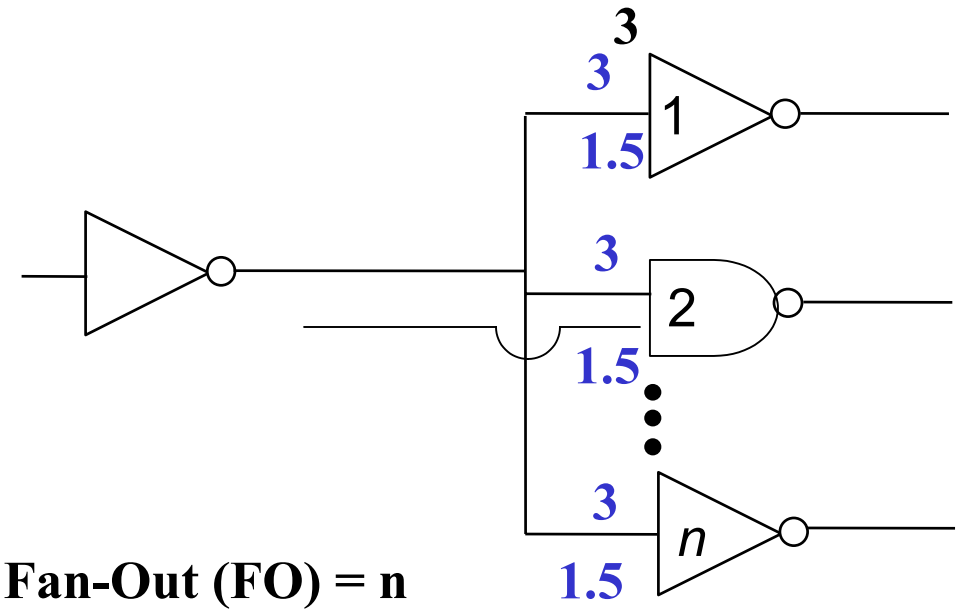
$\Delta t = 0.69(24k\Omega)(9\text{ fF}) = 149$ ps

For comparison the inverter had a pull-up delay of 30 ps

Fanout

Fanout is always ≥ 1 (there is always a load)

Gate capacitances sum and are charged by the driver resistance



Fan-Out (FO) = n

One load device was in included in the initial estimate of C_{LOAD}.

$$C'_{LOAD} = C_{LOAD} + (FO - 1)((W/L)_p + (W/L)_n)(C_{G/MS})$$

Assumes minimum length devices.

$$C'_{LOAD} = C_{LOAD} + (FO - 1)(1.5+3.0)(0.4\ fF)$$

Coping with Power Consumption

D.C. POWER **a.c. POWER**

Tube: 300V x 20 mA = 6W

Bipolar Transistor: 5V x 20 mA = 200 mW

NMOS Transistor: 5V x 200 μA = 1 mW

CMOS Transistors: 5V x 100 nA = 0.5 μW

True of every gate!

$$P_{SHORT\text{-}CIRCUIT} = (1/2)\, I_{SHORT\text{-}CIRCUIT}\, V_{DD}\, \tau_{30\text{-}70}\, f_{CLOCK}$$

$$= (1/2)\,(60\ \mu A)\,2.5V\,(0.1ns)\,(10^9) = 7.5\ \mu W$$

Assumes ½ of the gates change state

$$P_{DYNAMIC} = (1/2)(1/2)\, C\, V_{DD}^{\,2}\, f_{CLOCK}$$

$$= (1/2)\,(1/2)(10\ fF)\,(2.5)^2\,10^9 = 15.6\ \mu W$$

Only the L =>H takes energy from V_{DD}

True for only active gates.

What Might You be Buying for Christmas 2010?

Today's Technology: Pentium IV, 42 million transistors, 40-60 Watts, L = 0.18 μm

Intel's Statements in San Francisco Chronicle pp. E1, 11/26/01:

"We don't see any fundamental barrier (to Moore's Law of scaling).

Introduced a Tera-Hertz transistor for the 30 GHz (2010 generation)

There is a problem of 'leakage' of energy even when the transistors are in the off position.

Intel's new design stems the energy flow by using a different material for the existing insulator in one area of the transistor while adding another insulator on top of the silicon.

The claim the leakage is 10,000 times smaller.

"We're betting our entire business on this."

If Intel wins their bet in 2010 you can use your knowledge from EE 42 to explain to the sales person how this technology improvement works.