# CSE477
VLSI Digital Circuits
Fall 2002

Lecture 10: The Inverter, A Dynamic View

Mary Jane Irwin (www.cse.psu.edu/~mji)
www.cse.psu.edu/~cg477

[Adapted from Rabaey's *Digital Integrated Circuits*, ©2002, J. Rabaey et al.]

Inverter Propagation Delay

❑ Propagation delay is proportional to the time-constant of the network formed by the pull-down resistor and the load capacitance
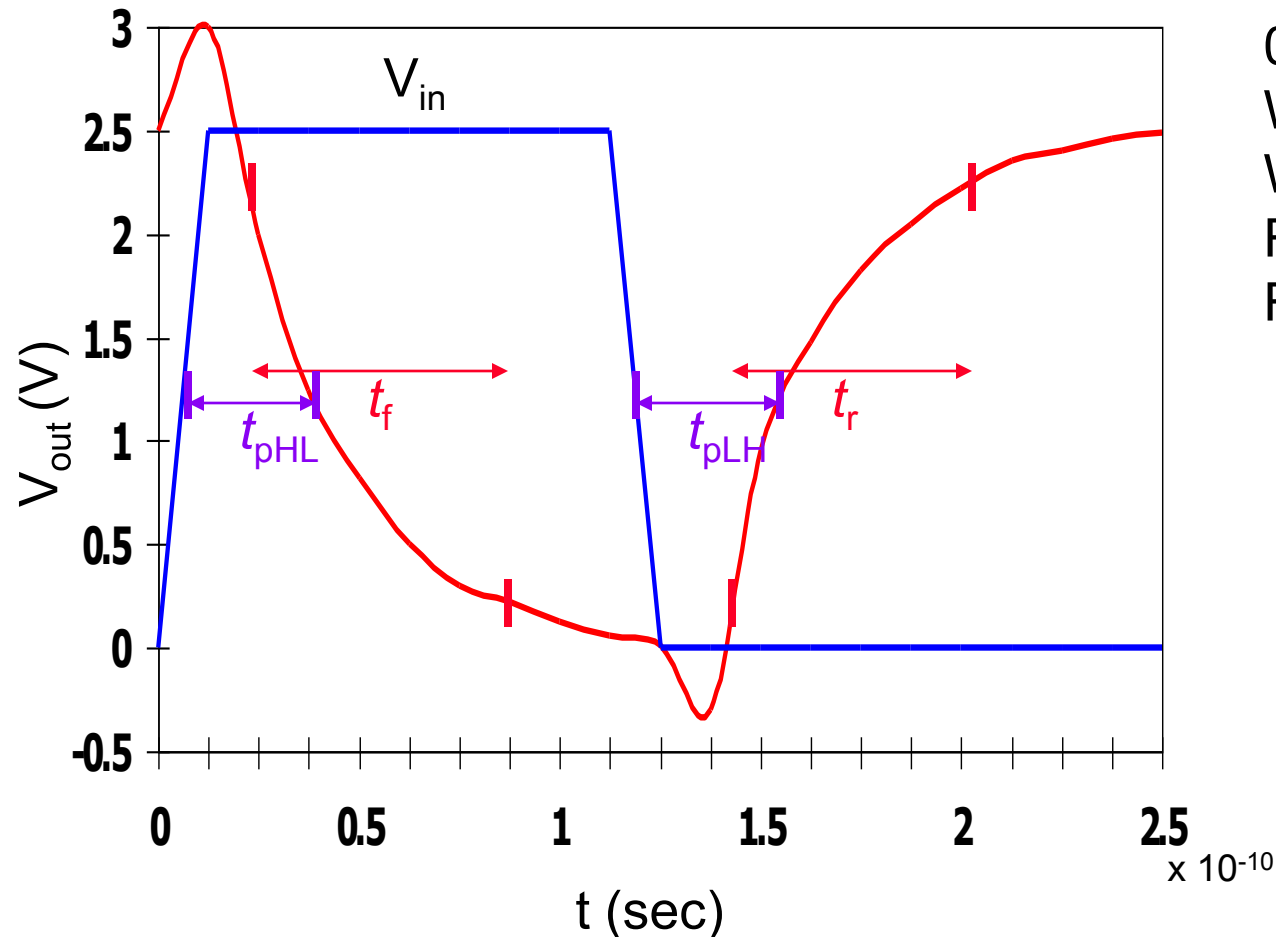
V_{DD}
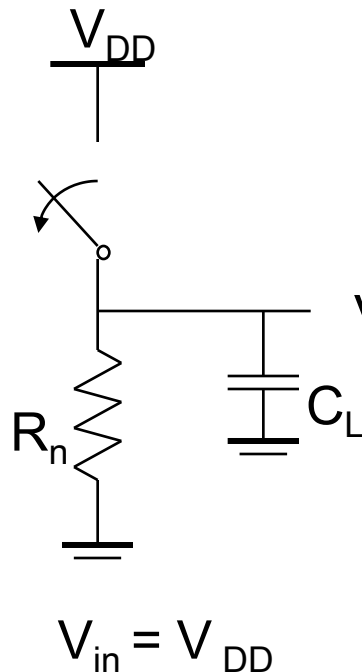
$$t_{pHL} = f(R_n, C_L)$$

$V_{out} = 0$

$$t_{pHL} = \ln(2)\ R_{eqn}\ C_L = 0.69\ R_{eqn}\ C_L$$

C_L

$$t_{pLH} = \ln(2)\ R_{eqp}\ C_L = 0.69\ R_{eqp}\ C_L$$

R_n

$V_{in} = V_{DD}$

$$t_p = (t_{pHL} + t_{pLH})/2 = 0.69\ C_L(R_{eqn} + R_{eqp})/2$$

❑ To equalize rise and fall times make the on-resistance of the NMOS and PMOS approximately equal.

Inverter Transient Response

V_{DD}=2.5V
0.25μm
W/L$_n$ = 1.5
W/L$_p$ = 4.5
R$_{eqn}$= 13 kΩ ($\div$ 1.5)
R$_{eqp}$= 31 kΩ ($\div$ 4.5)

t_{pHL} = 36 psec

t_{pLH} = 29 psec

so
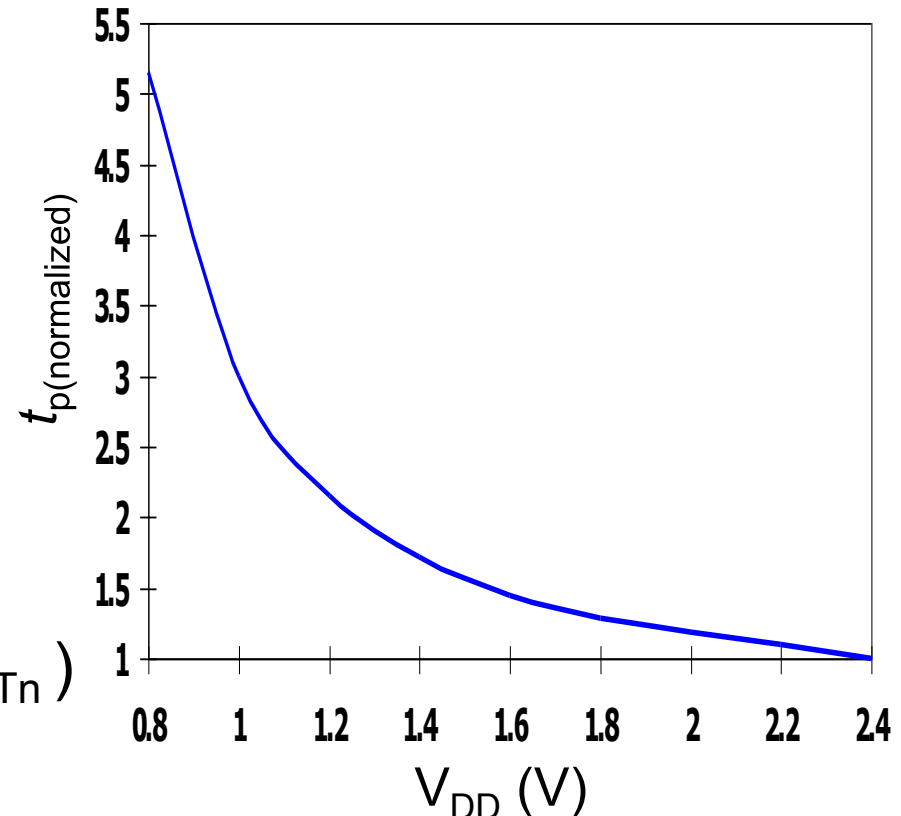
t_p = 32.5 psec

From simulation: t_{pHL} = 39.9 psec   and t_{pLH} = 31.7 psec

Inverter Propagation Delay, Revisited

❑ To see how a designer can optimize the delay of a gate have to expand the R_{eq} in the delay equation

$t_{pHL} = 0.69 \, R_{eqn} \, C_L$

$\quad = 0.69 \, (3/4 \, (C_L \, V_{DD})/I_{DSATn})$

$\quad \approx 0.52 \, C_L \, / \, (W/L_n \, k'_n \, V_{DSATn})$

Design for Performance

❑ Reduce C_L

  ☐ internal diffusion capacitance of the gate itself

    - keep the drain diffusion as small as possible

  ☐ interconnect capacitance

  ☐ fanout

❑ Increase W/L ratio of the transistor

  ☐ the most powerful and effective performance optimization tool in the hands of the designer

  ☐ watch out for self-loading! – when the intrinsic capacitance dominates the extrinsic load

❑ Increase V_{DD}

  ☐ can trade-off energy for performance

  ☐ increasing V_{DD} above a certain level yields only very minimal improvements

 ☐ reliability concerns enforce a firm upper bound on V_{DD}

NMOS/PMOS Ratio

❑ So far have sized the PMOS and NMOS so that the R_{eq}'s match (ratio of 3 to 3.5)

  ❑ symmetrical VTC

  ❑ equal high-to-low and low-to-high propagation delays

❑ If speed is the only concern, reduce the width of the PMOS device!

  ❑ widening the PMOS degrades the t_{pHL} due to larger parasitic capacitance

$$\beta = (W/L_p)/(W/L_n)$$

$r = R_{eqp}/R_{eqn}$ (resistance ratio of identically-sized PMOS and NMOS)

$$\beta_{opt} = \sqrt{r} \text{ when wiring capacitance is negligible}$$

PMOS/NMOS Ratio Effects
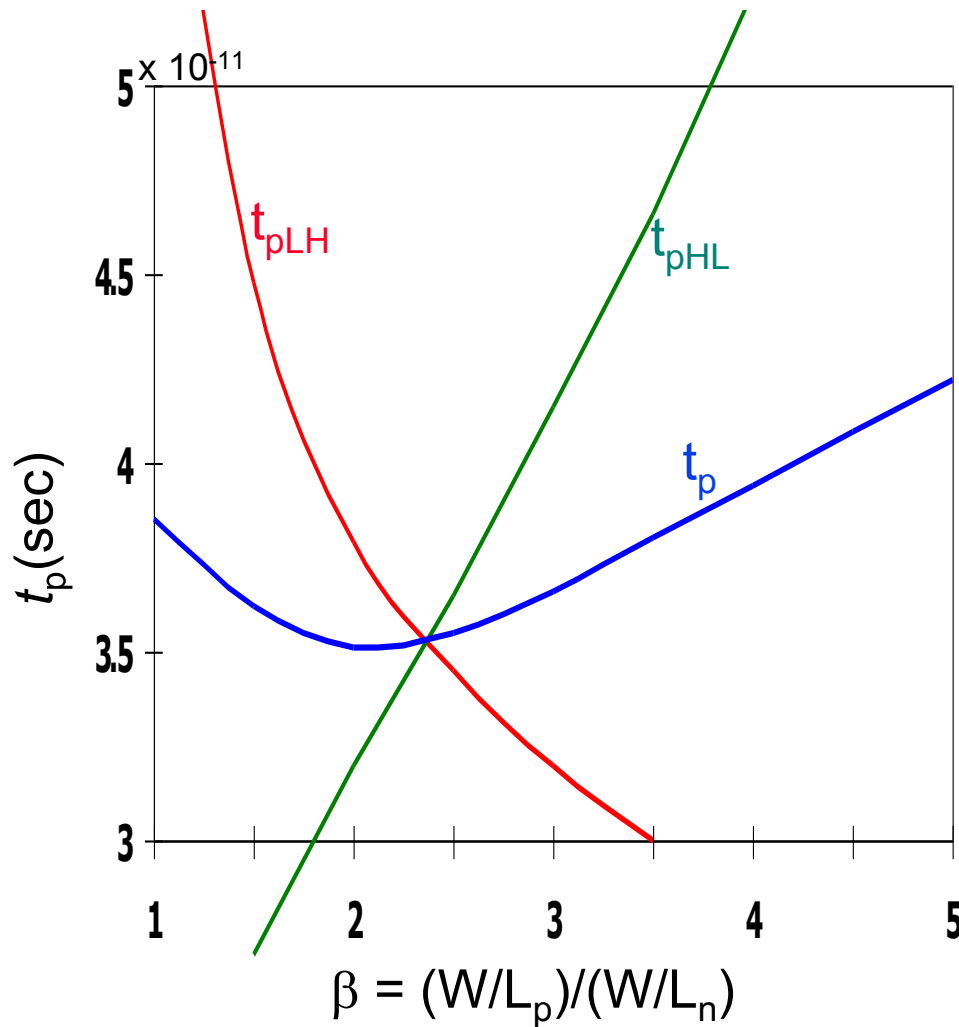
β of 2.4 (= 31 kΩ/13 kΩ) gives symmetrical response

β of 1.6 to 1.9 gives optimal performance

Device Sizing for Performance

❑ Divide capacitive load, C_L, into

 ▢ C_{int} : intrinsic - diffusion and Miller effect

 ▢ C_{ext} : extrinsic - wiring and fanout

$$t_p = 0.69\ R_{eq}\ C_{int}\ (1 + C_{ext}/C_{int}) = t_{p0}\ (1 + C_{ext}/C_{int})$$

 ▢ where $t_{p0} = 0.69\ R_{eq}\ C_{int}$ is the intrinsic (unloaded) delay of the gate

❑ Widening both PMOS and NMOS by a factor S reduces R_{eq} by an identical factor ($R_{eq} = R_{ref}/S$), but raises the intrinsic capacitance by the same factor ($C_{int} = SC_{iref}$)

$$t_p = 0.69\ R_{ref}\ C_{iref}\ (1 + C_{ext}/(SC_{iref})) = t_{p0}(1 + C_{ext}/(SC_{iref}))$$

 ▢ t_{p0} is independent of the sizing of the gate; *with no load the drive of the gate is totally offset by the increased capacitance*

 ▢ any S sufficiently larger than (C_{ext}/C_{int}) yields the best performance gains with least area impact

Sizing Impacts on Delay

for a fixed load

t_p(sec), $\times 10^{-11}$, S axis from 1 to 15
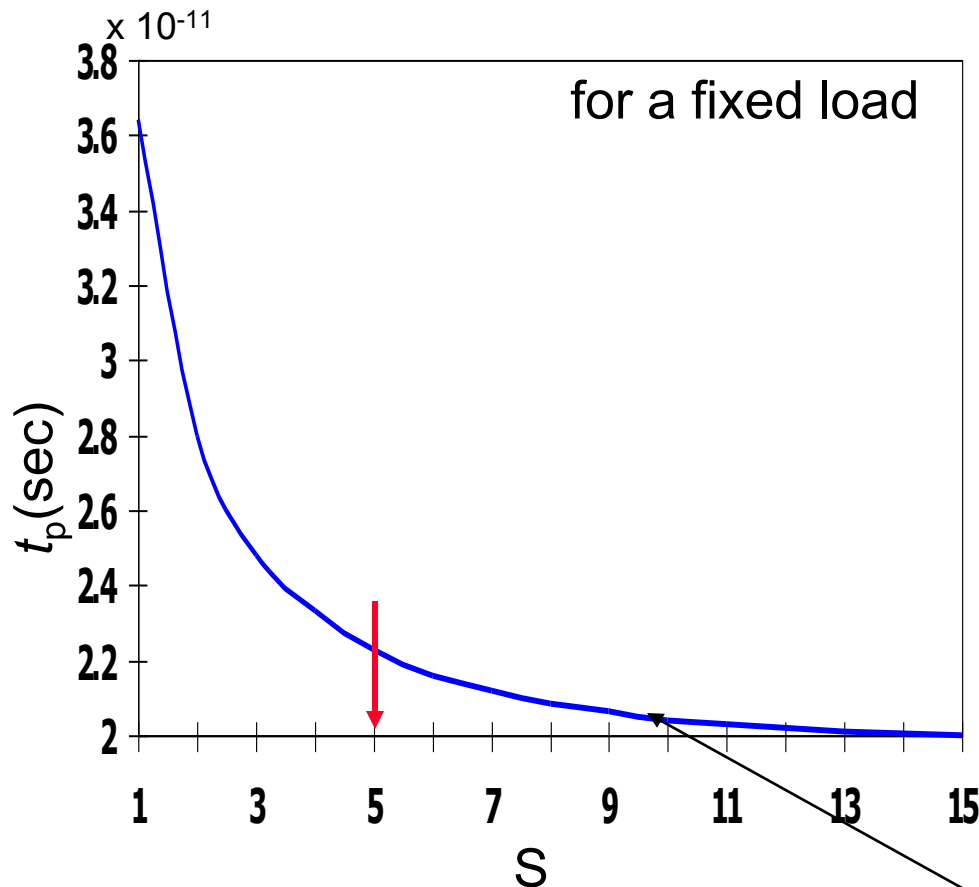
The majority of the improvement is already obtained for S = 5. Sizing factors larger than 10 barely yield any extra gain (and cost significantly more area).

self-loading effect (intrinsic capacitance dominates)

Impact of Fanout on Delay

❑ Extrinsic capacitance, C_{ext}, is a function of the fanout of the gate - the larger the fanout, the larger the external load.

❑ First determine the input loading effect of the inverter. Both C_g and C_{int} are proportional to the gate sizing, so $C_{int} = \gamma C_g$ is independent of gate sizing and
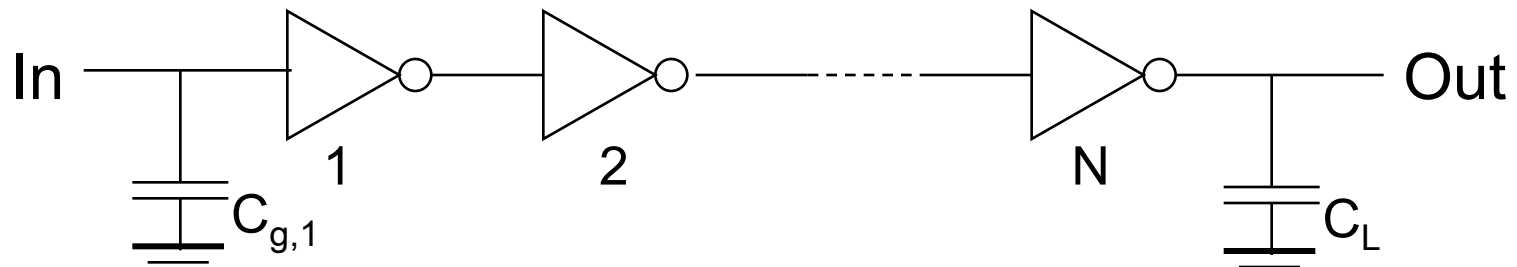
$$t_p = t_{p0} (1 + C_{ext}/ \gamma C_g) = t_{p0} (1 + f/\gamma)$$

i.e., the delay of an inverter is a function of the ratio between its external load capacitance and its input gate capacitance: the effective fan-out f

$$f = C_{ext}/C_g$$

Inverter Chain

❑ Real goal is to minimize the delay through an inverter chain

the delay of the j-th inverter stage is

$$t_{p,j} = t_{p0} \left(1 + C_{g,j+1}/(\gamma C_{g,j})\right) = t_{p0}(1 + f_j/\gamma)$$

and $\qquad t_p = t_{p1} + t_{p2} + \ldots + t_{pN}$

so $\qquad t_p = \sum t_{p,j} = t_{p0} \sum \left(1 + C_{g,j+1}/(\gamma C_{g,j})\right)$

❑ If C_L is given

  ☐ How should the inverters be sized?

  ☐ How many stages are needed to minimize the delay?

Sizing the Inverters in the Chain

❑ The optimum size of each inverter is the geometric mean of its neighbors – meaning that if each inverter is sized up by the same factor f wrt the preceding gate, it will have the same effective fan-out and the same delay

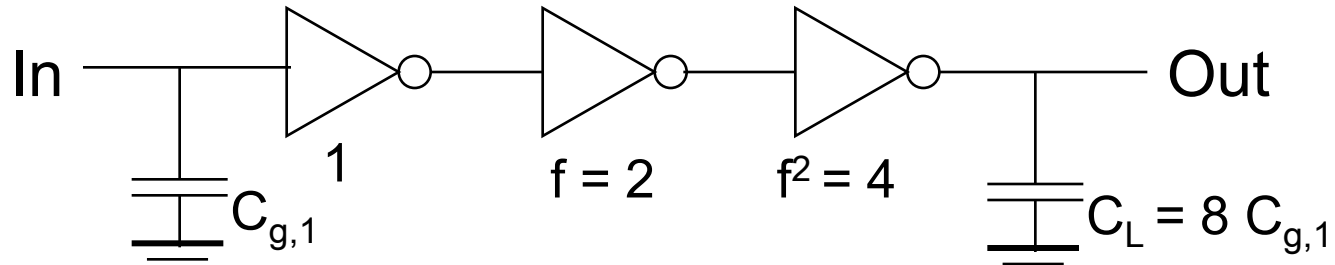$$f = \sqrt[N]{C_L/C_{g,1}} = \sqrt[N]{F}$$

where F represents the overall effective fan-out of the circuit ($F = C_L/C_{g,1}$)

and the minimum delay through the inverter chain is

$$t_p = N\, t_{p0} \left(1 + (\sqrt[N]{F})/\gamma\right)$$

❑ The relationship between t_p and F is linear for one inverter, square root for two, etc.

Example of Inverter Chain Sizing

❑ $C_L/C_{g,1}$ has to be evenly distributed over N = 3 inverters

$$C_L/C_{g,1} = 8/1$$

$$f = \sqrt[3]{8} = 2$$
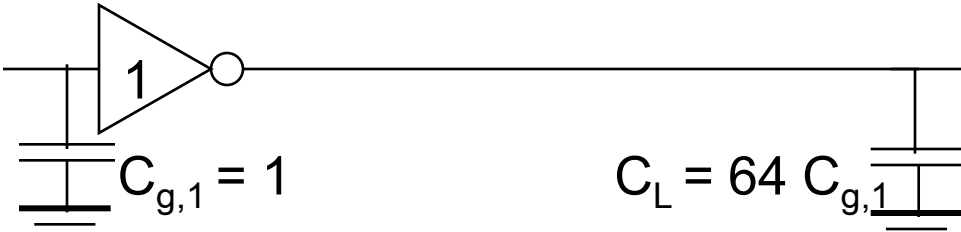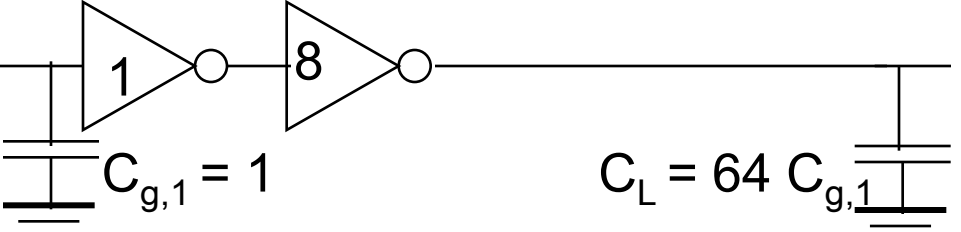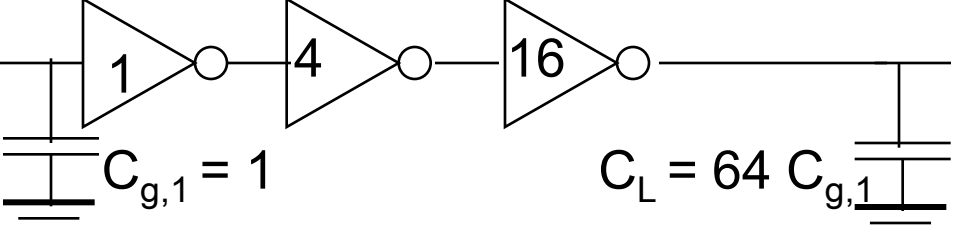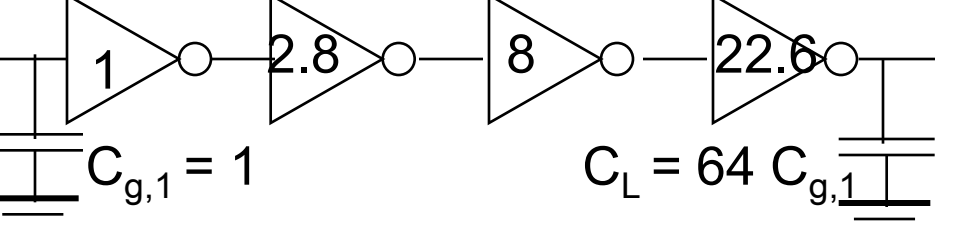
Determining N: Optimal Number of Inverters

- What is the optimal value for N given F ($=f^N$) ?

  - if the number of stages is too large, the intrinsic delay of the stages becomes dominate

  - if the number of stages is too small, the effective fan-out of each stage becomes dominate

- The optimum N is found by differentiating the minimum delay expression divided by the number of stages and setting the result to 0, giving

$$\gamma + \sqrt[N]{F} - (\sqrt[N]{F} \, \ln F)/N = 0$$

- For $\gamma = 0$ (ignoring self-loading) N = ln (F) and the effective-fan out becomes f = e = 2.71828

- For $\gamma = 1$ (the typical case) the optimum effective fan-out (tapering factor) turns out to be close to 3.6

Optimum Effective Fan-Out
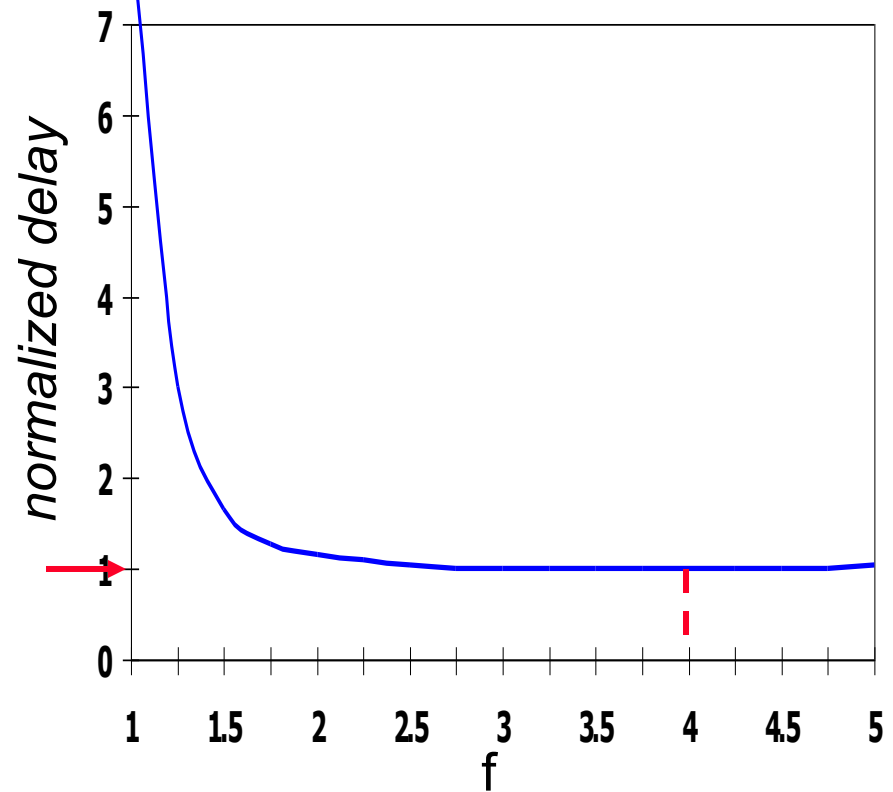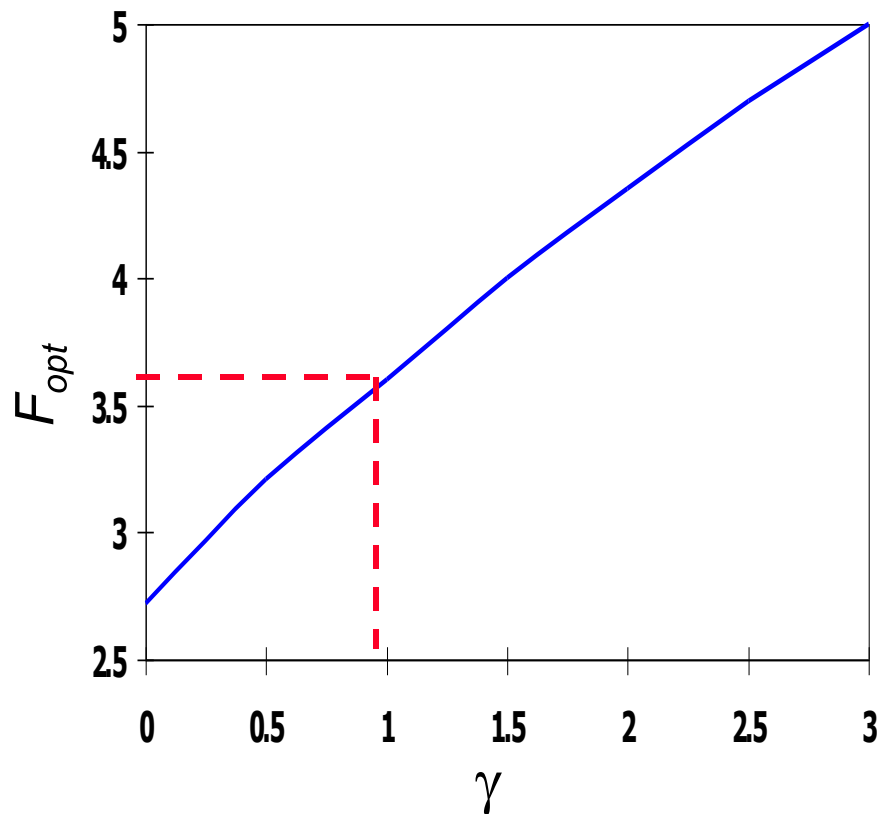
❑ Choosing f larger than optimum has little effect on delay and reduces the number of stages (and area).

    ▫ Common practice to use f = 4 (for γ = 1)

    ▫ But too many stages has a substantial negative impact on delay

Example of Inverter (Buffer) Staging

N	f	t_p
1	64	65
2	8	18
3	4	15
4	2.8	15.3

Impact of Buffer Staging for Large C_L

F ($\gamma = 1$)	Unbuffered	Two Stage Chain	Opt. Inverter Chain
10	11	8.3	8.3
100	101	22	16.5
1,000	1001	65	24.8
10,000	10,001	202	33.1

❑ Impressive speed-ups with optimized cascaded inverter chain for very large capacitive loads.

Input Signal Rise/Fall Time

❑ In reality, the input signal changes gradually (and both PMOS and NMOS conduct for a brief time). This affects the current available for charging/discharging C_L and impacts propagation delay.

❑ t_p increases linearly with increasing input slope, t_s, once $t_s > t_p$

❑ t_s is due to the limited driving capability of the preceding gate

x 10⁻¹¹

(graph: y-axis labeled t_p(sec) ranging from 3.6 to 5.4, x-axis labeled t_s(sec) ranging from 0 to 8 x 10⁻¹¹, with an upward curve)
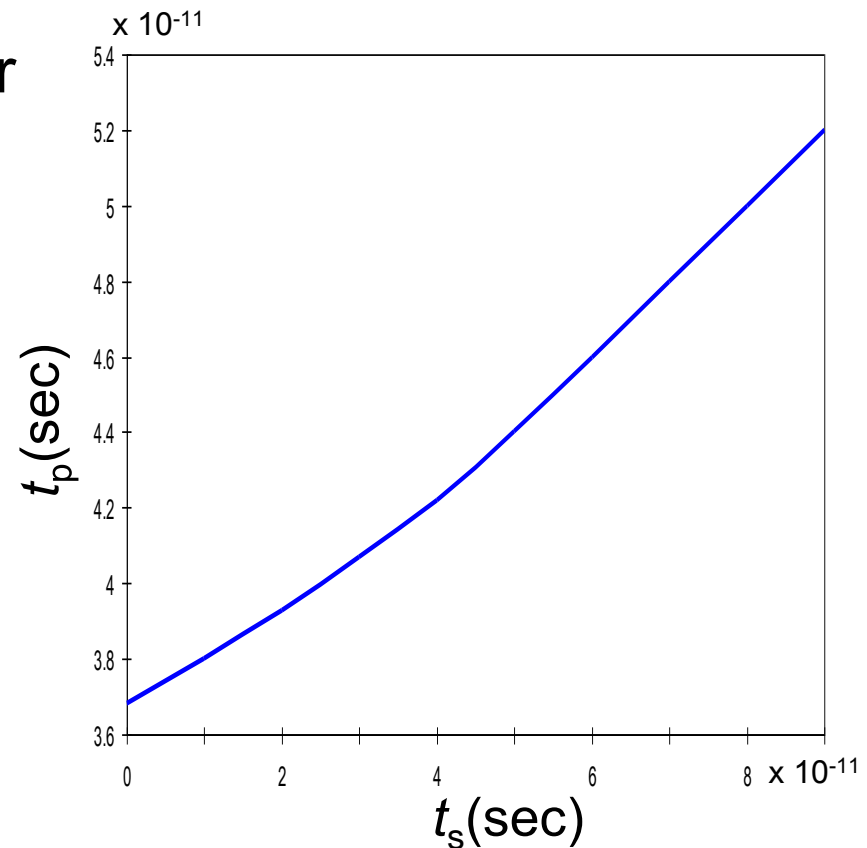
for a minimum-size inverter with a fan-out of a single gate

Design Challenge

❑ A gate is never designed in isolation: its performance is affected by both the fan-out and the driving strength of the gate(s) feeding its inputs.

$$t^i_p = t^i_{step} + \eta \, t^{i-1}_{step} \qquad (\eta \approx 0.25)$$

❑ Keep signal rise times smaller than or equal to the gate propagation delays.

   ❑ good for performance

   ❑ good for power consumption

❑ Keeping rise and fall times of the signals small and of approximately equal values is one of the major challenges in high-performance designs - slope engineering.

Delay with Long Interconnects

❑ When gates are farther apart, wire capacitance and resistance can no longer be ignored.
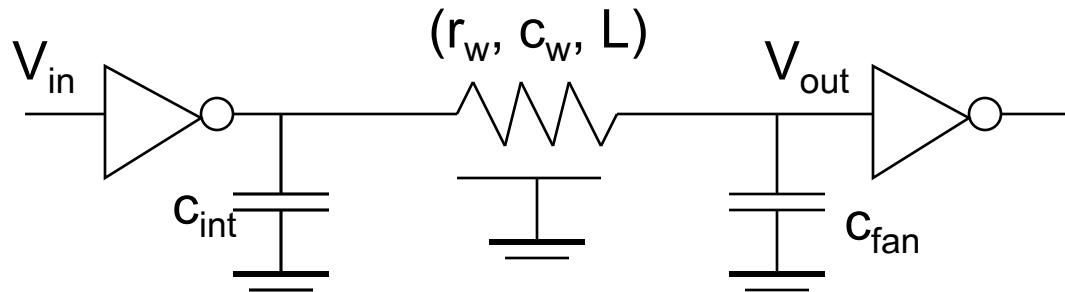
$$t_p = 0.69R_{dr}C_{int} + (0.69R_{dr}+0.38R_w)C_w + 0.69(R_{dr}+R_w)C_{fan}$$

where $R_{dr} = (R_{eqn} + R_{eqp})/2$

$$= 0.69R_{dr}(C_{int}+C_{fan}) + 0.69(R_{dr}c_w+r_wC_{fan})L + 0.38r_wc_wL^2$$

❑ Wire delay rapidly becomes the dominate factor (due to the quadratic term) in the delay budget for longer wires.

Next Lecture and Reminders

- ❑ Next lecture
  - ᴄ Designing fast logic
    - Reading assignment – Rabaey, et al, 6.2.1

- ❑ Reminders
  - ᴄ Project specifications due today
  - ᴄ HW3 due next Thursday, Oct 10th (hand in to TA)
  - ᴄ Class cancelled on Oct 10th as make up for evening midterm
  - ᴄ I will be out of town Oct 10th through Oct 15th and Oct 18th through Oct 23rd, so office hours during those periods are cancelled
  - ᴄ We will have a guest lecturer on Oct 22nd
  - ᴄ Evening midterm exam scheduled
    - Wednesday, October 16th from 8:15 to 10:15pm in 260 Willard
    - Only one midterm conflict filed for so far