

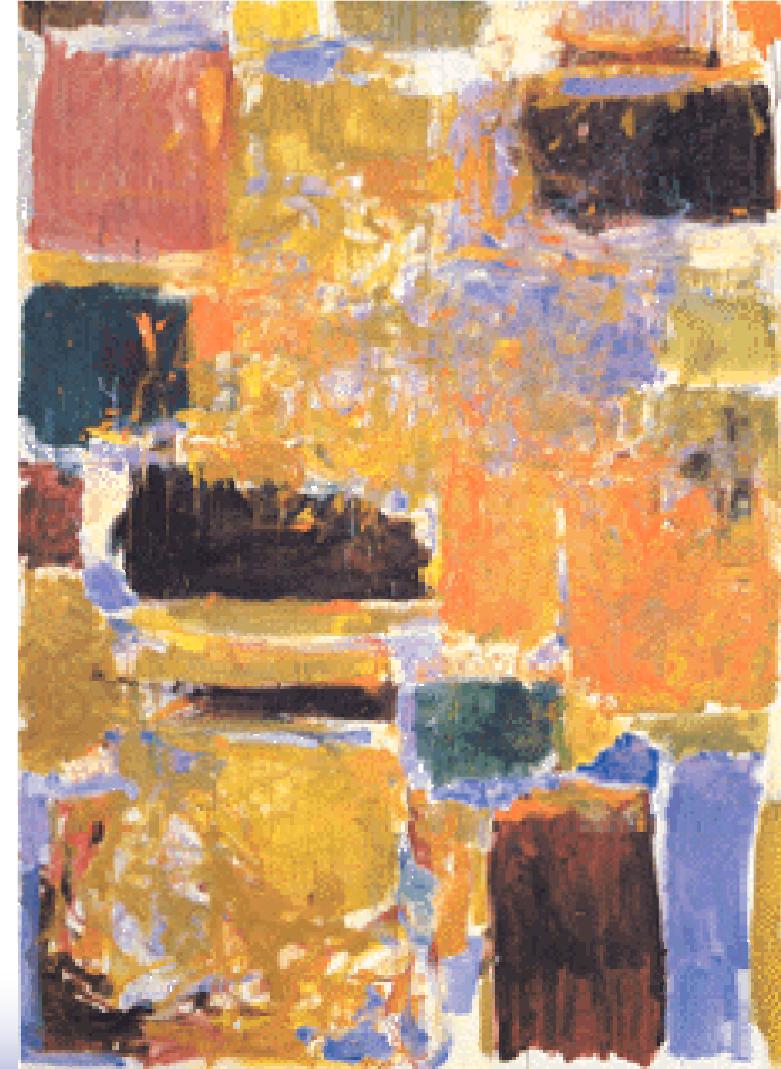


CS150 - EE141/241A
Fall 2014
Digital Design and
Integrated Circuits

Instructors:
John Wawrzynek and Vladimir Stojanovic

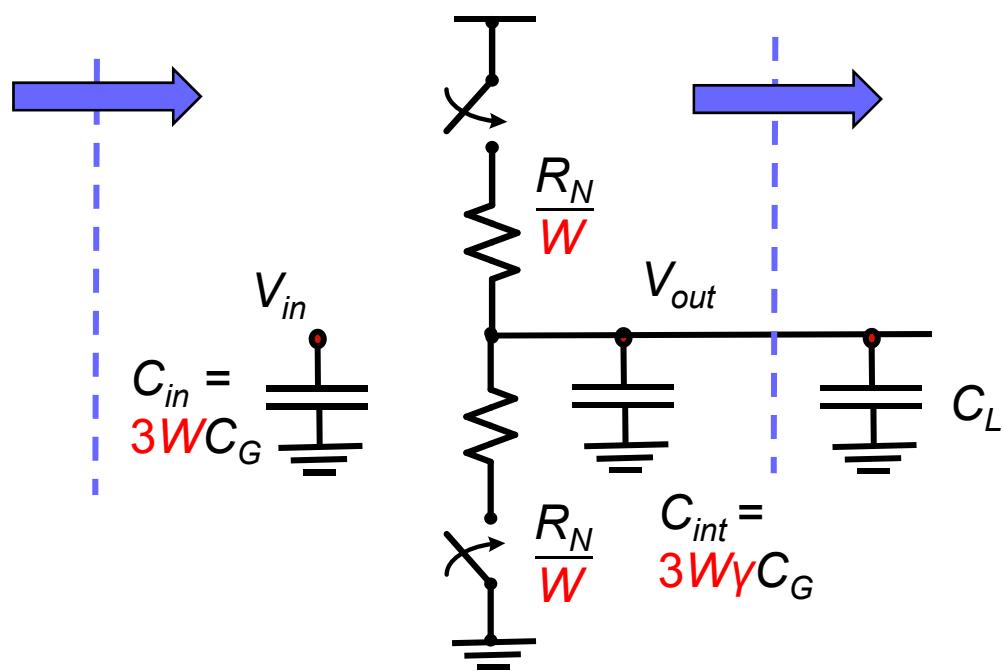
Lecture 9

Outline



- Inverter Delay Optimization
- Logical Effort
- Optimizing Complex Combinational Logic

Inverter with Load Capacitance



$$\begin{aligned}
 t_p &= 0.69 \left(\frac{R_N}{W} \right) (C_{int} + C_L) \\
 &= 0.69 \left(\frac{R_N}{W} \right) (3W\gamma C_G + C_L) \\
 &= 0.69 (3C_G R_N) \left(\gamma + \frac{C_L}{C_{in}} \right) \\
 &= \underline{t_{inv}} (\gamma + \frac{C_L}{C_{in}}) = \underline{t_0} (\gamma + f)
 \end{aligned}$$

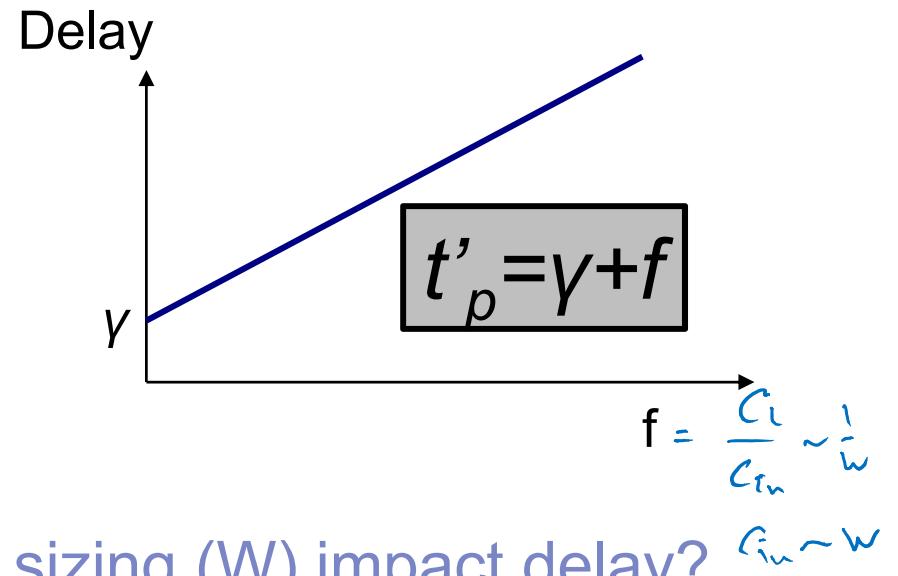
$f = \text{fanout} = \text{ratio between load and input capacitance of gate}$

Inverter Delay Model

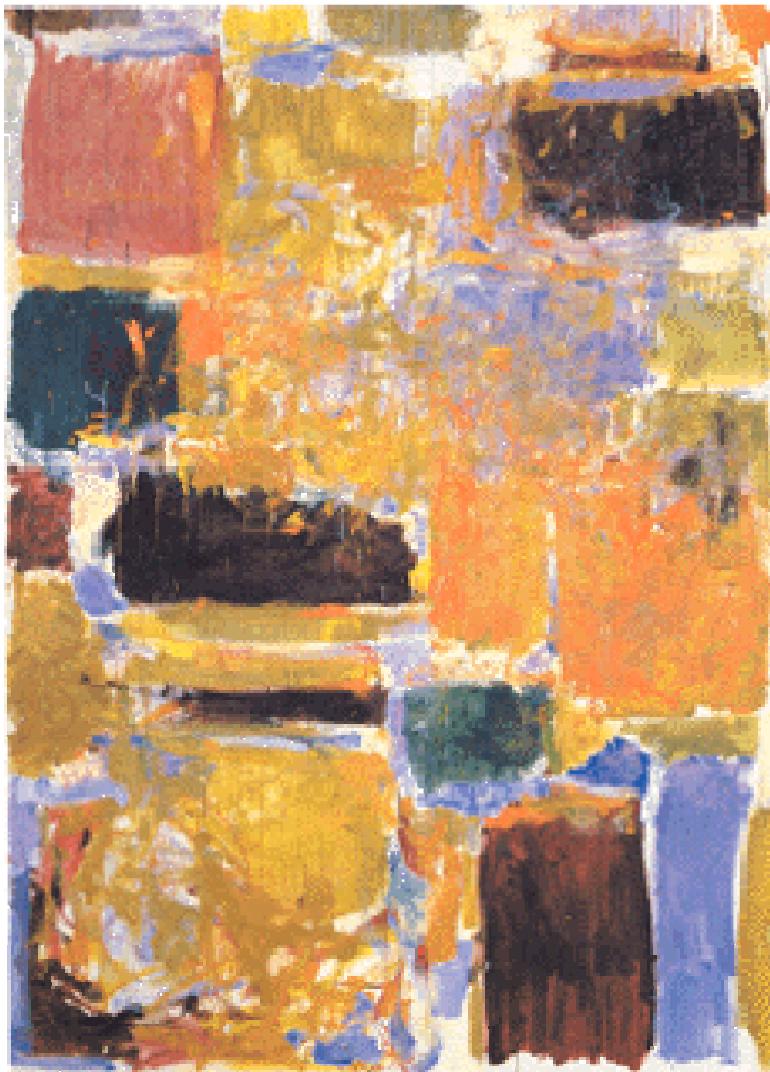
$$t_p = t_{inv}(\gamma + f)$$

t_{inv} technology constant

- Can be dropped from expression
- Delay unit-less variable (expressed in unit delays)

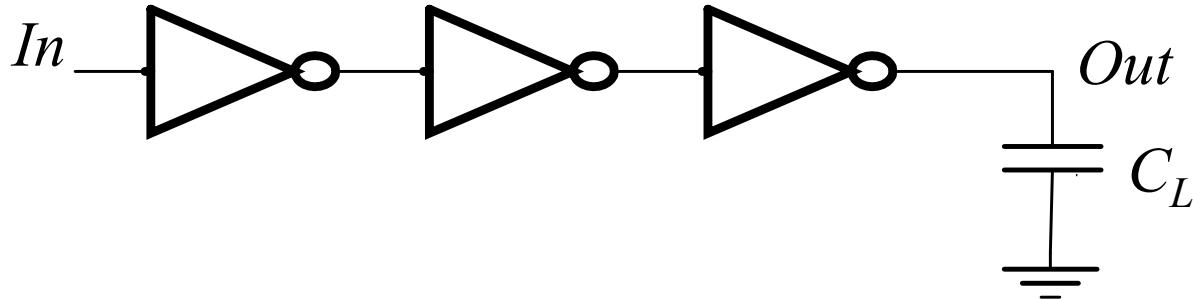


Question: how does transistor sizing (W) impact delay?



Inverter Delay Optimization

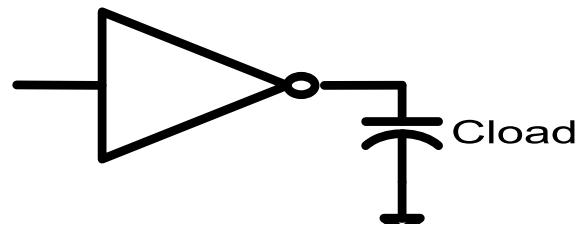
Inverter Chain



- For some given C_L :
 - How many stages are needed to minimize delay?
 - How to size the inverters?
- Anyone want to guess the solution?

Careful about Optimization Problems

- Get fastest delay if build one **very** big inverter
 - So big that delay is set only by self-loading



- Likely not the problem you're interested in
 - Someone has to drive this inverter...

Engineering Optimization Problems in General

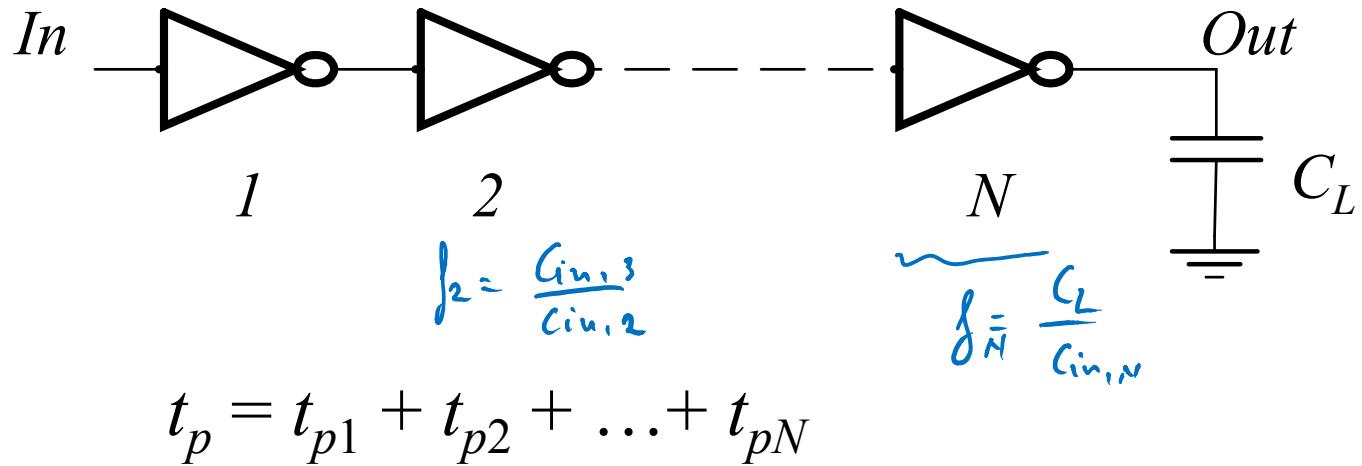
- Need to have a set of constraints
- Constraints key to:
 - Making the result useful
 - Making the problem have a ‘clean’ solution
- For sizing problem:
 - Need to constrain size of first inverter

Delay Optimization Problem #1

- You are given:
 - A fixed number of inverters
 - The size of the first inverter
 - The size of the load that needs to be driven
- Your goal:
 - Minimize the delay of the inverter chain
- Need model for inverter delay vs. size

$$t_p = t_{inv} (\gamma + \delta)$$

Apply to Inverter Chain



$$t_{p_j} = t_{inv} \left(\gamma + \frac{C_{in,j+1}}{C_{in,j}} \right)$$

$$t_p = \sum_{j=1}^N t_{p,j} = t_{inv} \sum_{i=1}^N \left(\gamma + \frac{C_{in,j+1}}{C_{in,j}} \right), \quad C_{in,N+1} = C_L$$

Optimal Sizing for Given N

- Delay equation has $N-1$ unknowns, $C_{in,2} \dots C_{in,N}$
- **To minimize the delay,** find $N-1$ partial derivatives:

$$t_p = \dots + t_{inv} \frac{C_{in,j}}{C_{in,j-1}} + t_{inv} \frac{C_{in,j+1}}{C_{in,j}} + \dots$$

$$\frac{dt_p}{dC_{in,j}} = t_{inv} \frac{1}{C_{in,j-1}} - t_{inv} \frac{C_{in,j+1}}{C_{in,j}^2} = 0$$

Optimal Sizing for Given N (cont'd)

- Result: every stage has equal fanout (f):

$$\frac{C_{in,j}}{C_{in,j-1}} = \frac{C_{in,j+1}}{C_{in,j}}$$

- Size of each stage is geometric mean of two neighbors:

$$C_{in,j} = \sqrt{C_{in,j-1} C_{in,j+1}}$$

- Equal fanout → every stage will have same delay

Optimum Delay and Number of Stages

- When each stage has same fanout f :

$$f^N = F = C_L / C_{in,1}$$

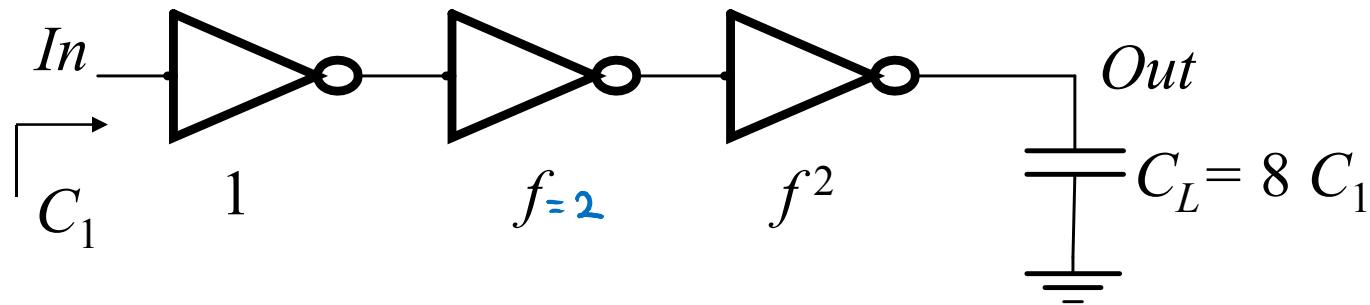
- Fanout of each stage:

$$f = \sqrt[N]{F}$$

- Minimum path delay:

$$t_p = N t_{inv} \left(\gamma + \sqrt[N]{F} \right)$$

Example



C_L/C_1 has to be evenly distributed across $N=3$ stages:

$$F = 8 \quad \Rightarrow \quad f = \sqrt[3]{8} = 2$$

Delay Optimization Problem #2

- You are given:
 - The size of the first inverter
 - The size of the load that needs to be driven
- Your goal:
 - Minimize delay by finding optimal number and sizes of gates
- So, need to find N that minimizes:

$$t_p = N t_{inv} \left(\gamma + \sqrt[N]{C_L / C_{in}} \right)$$

Untangling the Optimization Problem

- Rewrite N in terms of fanout/stage f :

$$f^N = C_L / C_{in} \rightarrow N = \frac{\ln(C_L / C_{in})}{\ln f}$$

$$t_p = N t_{inv} \left((C_L / C_{in})^{1/N} + \gamma \right) = t_{inv} \ln(C_L / C_{in}) \left(\frac{f + \gamma}{\ln f} \right)$$

$$\frac{\partial t_p}{\partial f} = t_{inv} \ln(C_L / C_{in}) \cdot \frac{\ln f - 1 - \gamma/f}{\ln^2 f} = 0$$

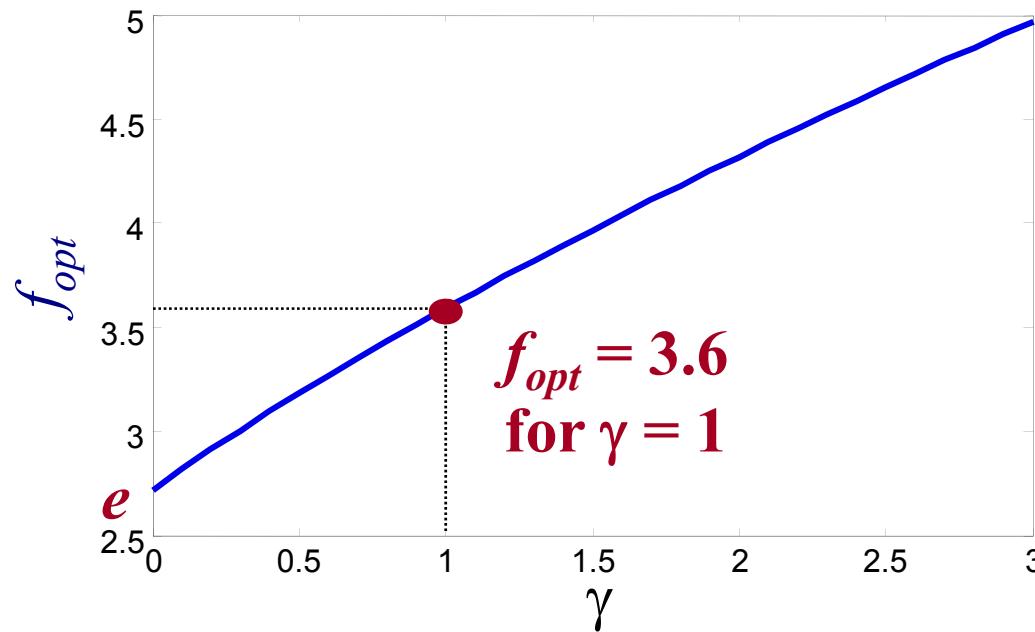
$$f = \exp(1 + \gamma/f) \quad (\text{no explicit solution})$$

For $\gamma=0, f=e, N=\ln(C_L/C_{in})$

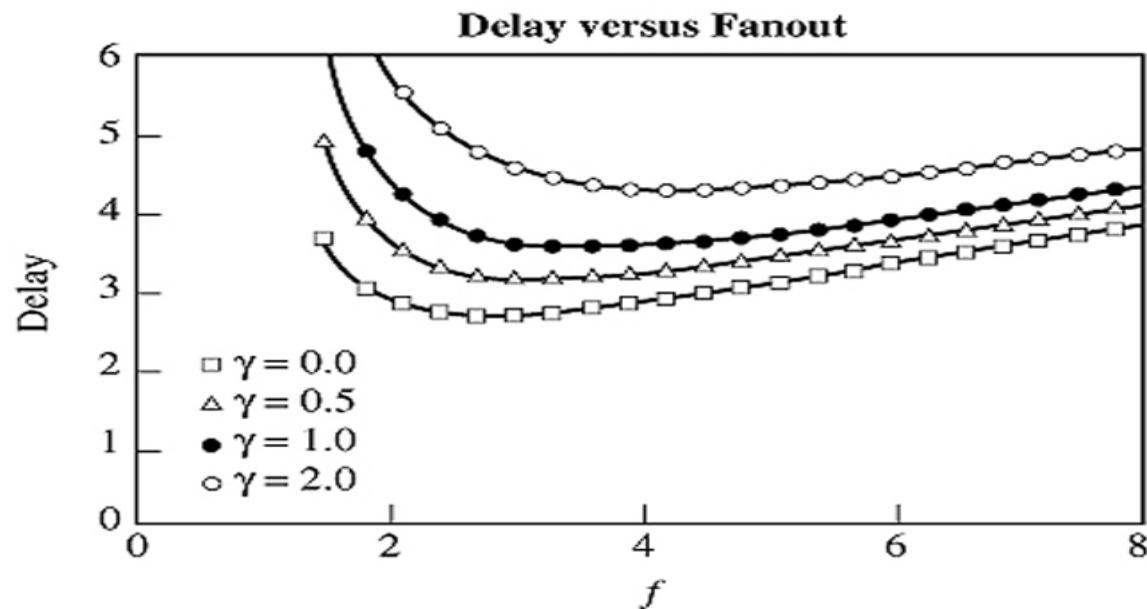
Optimum Effective Fanout f

- Optimum f for given process defined by γ

$$f = \exp(1 + \gamma / f)$$



In Practice: Plot of Total Delay



[Hodges, p.281]

- Curves very flat for $f > 2$
 - Simplest/most common choice: $f = 4$

Normalized Delay As a Function of F

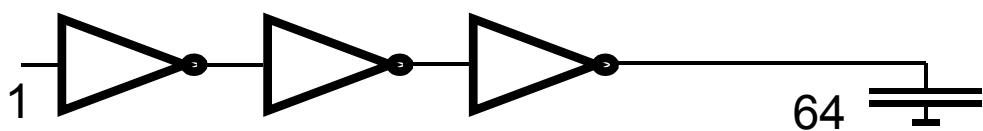
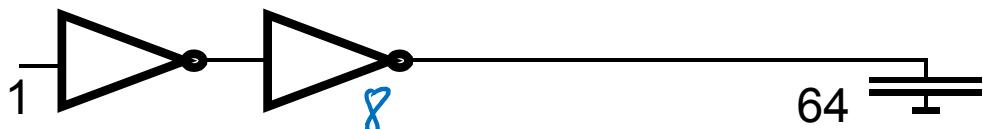
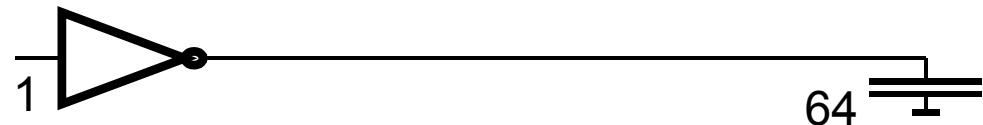
$$t_p = N t_{inv} \left(\gamma + \sqrt[N]{F} \right), \quad F = C_L / C_{in}$$

F	Unbuffered	Two Stage	Inverter Chain
10	11	8.3	8.3
100	101	22	16.5
1000	1001	65	24.8
10,000	10,001	202	33.1

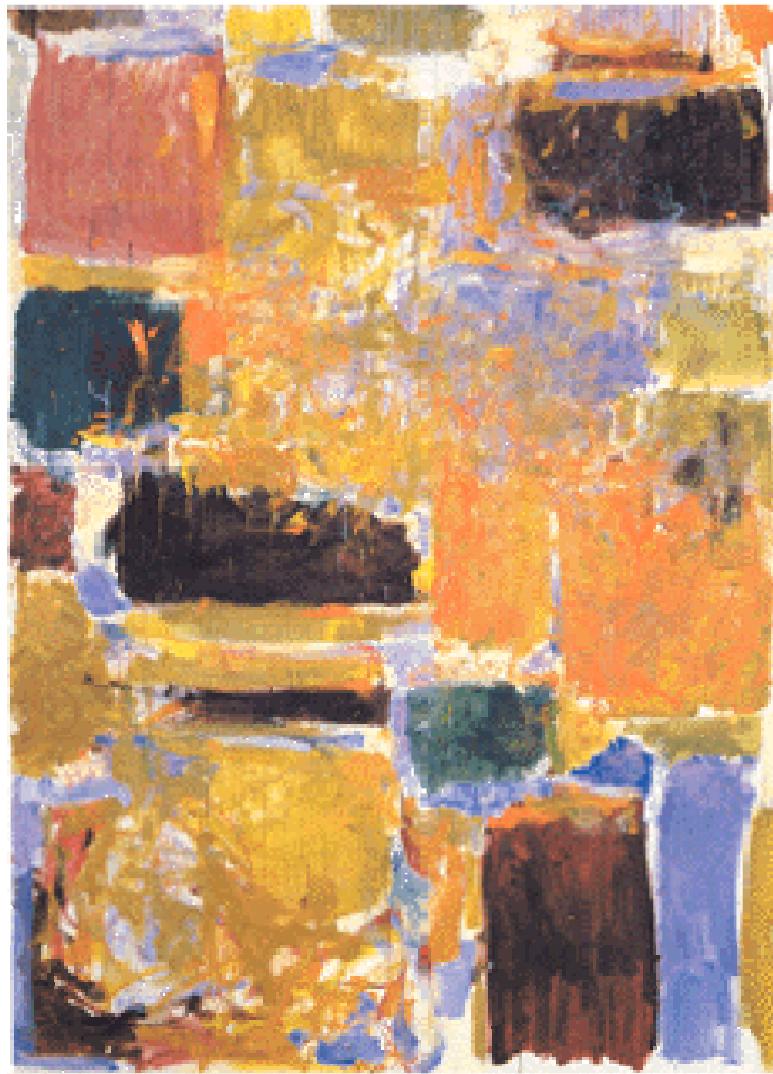
$$(\gamma = 1)$$

[Rabaey: page 210]

Buffer Design



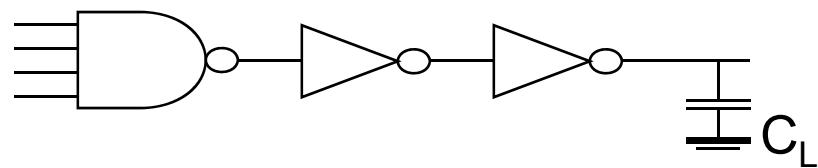
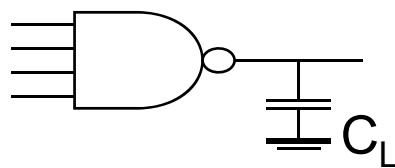
N	f	t_p
1	64	65
2	8	$2 \cdot 8 + 2 = 18$
3	4	$3 \cdot 4 + 3 = 15$



Logical Effort

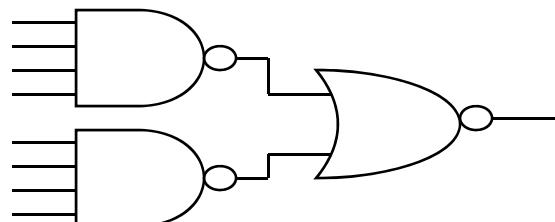
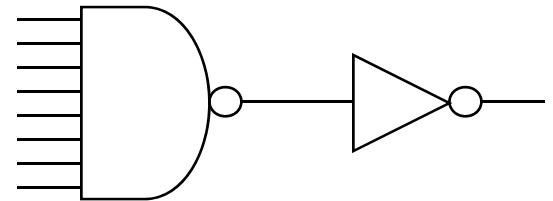
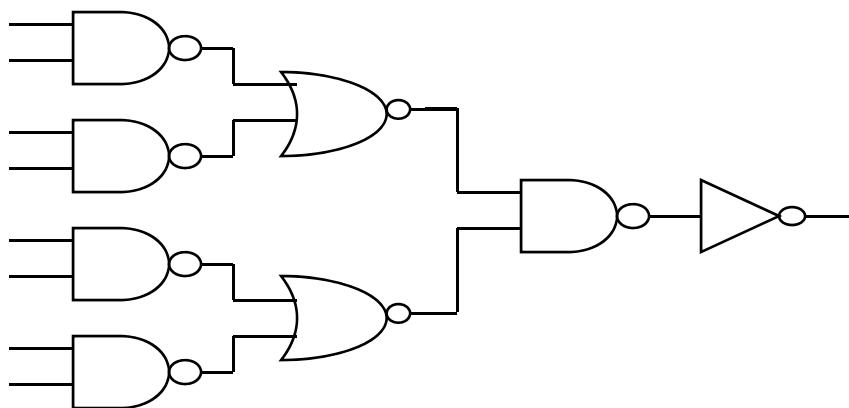
Question #1

- How to best combine logic and drive for a big capacitive load?



Question #2

- All of these are “decoders”
 - Which one is “best”?



Method to answer both of these questions

- Extension of buffer sizing problem
- Logical effort

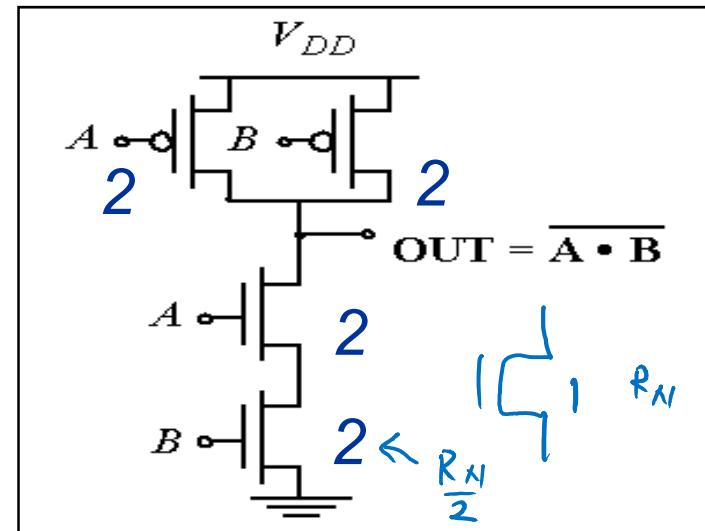
Complex Gate Sizing: NAND-2 Example

$$C_{gnand} = 4C_G = (4/3) C_{ginv}$$

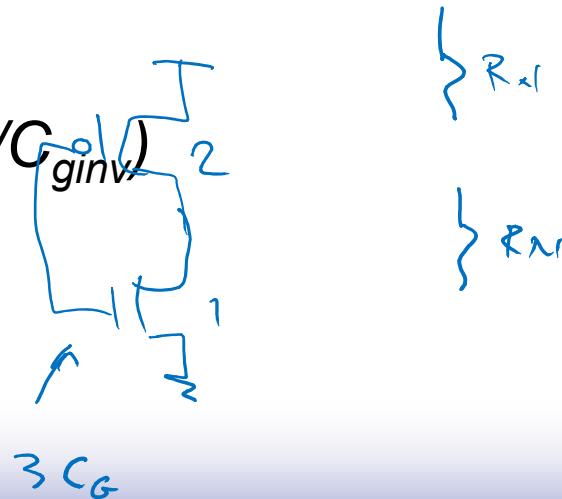
$$C_{dnand} = 6C_D = 6\gamma C_G = 2\gamma C_{ginv}$$

$$f = C_L/C_{gnand} = (3/4) C_L/C_{ginv}$$

$$C_{in, A, B} = 4C_G$$



$$\begin{aligned} t_{pNAND} &= kR_N(C_{dnand} + C_L) \\ &= kR_N(2\gamma C_{ginv} + C_L) \\ &= kR_N C_{ginv} (2\gamma + C_L/C_{ginv}) \\ &= t_{inv} (2\gamma + (4/3)f) \end{aligned}$$



Logical Effort

- Defines ease of gate to drive external capacitance
- Inverter has the smallest logical effort and intrinsic delay of all static CMOS gates
- Logical effort LE is defined as:
 - $(R_{eq,gate} C_{in,gate}) / (R_{eq,inv} C_{in,inv})$
 - Easiest way to calculate (usually):
 - Size gate to deliver same current as an inverter, take ratio of gate input capacitance to inverter capacitance
- LE increases with gate complexity

Logical Effort

$$t_{pgate} = t_{inv} (p + LEf)$$

Measure everything in units of t_{inv} (divide by t_{inv}):

p – intrinsic delay - gate parameter $\neq f(W)$

LE – logical effort – gate parameter $\neq f(W)$

f – electrical fanout = $C_L/C_{in} = f(W) \sim \frac{1}{W}$

Normalize everything to an inverter:

$$LE_{inv} = 1, p_{inv} = \gamma$$

Delay of a Logic Gate

Gate delay:

$$\text{Delay} = EF + p \quad (\text{measured in units of } t_{inv})$$

effective fanout **intrinsic delay**

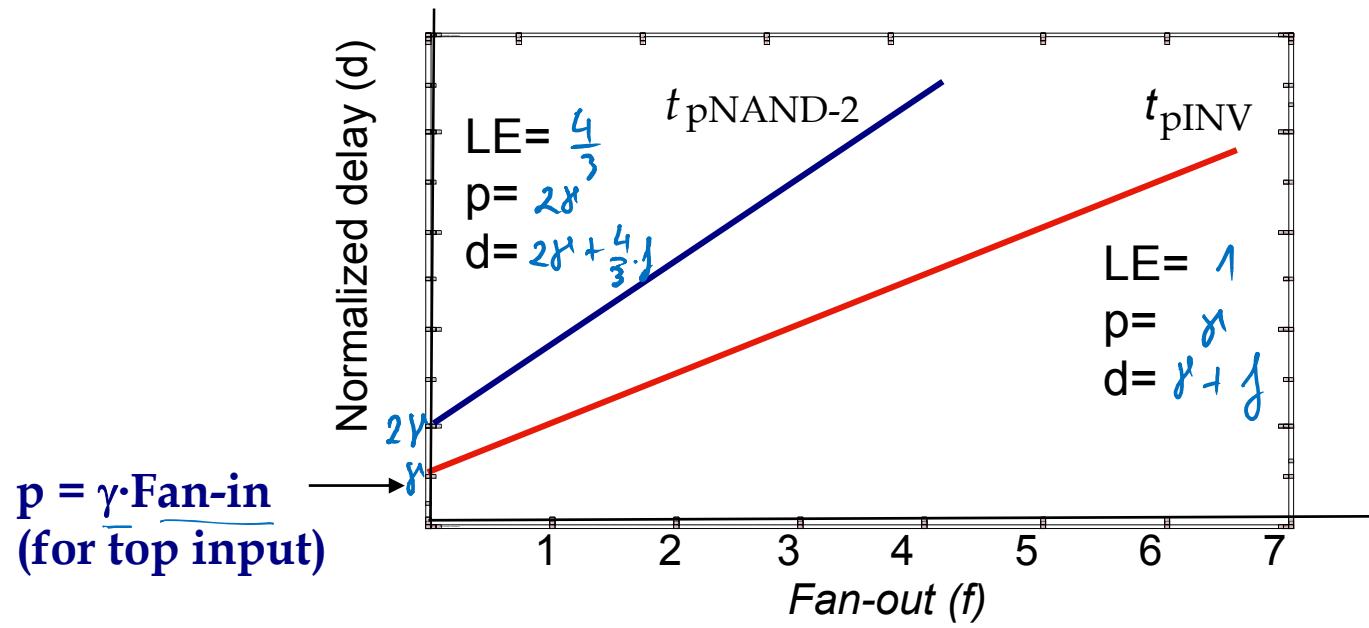
Effective fanout:

$$EF = LE f$$

logical effort **electrical fanout** = C_L/C_{in}

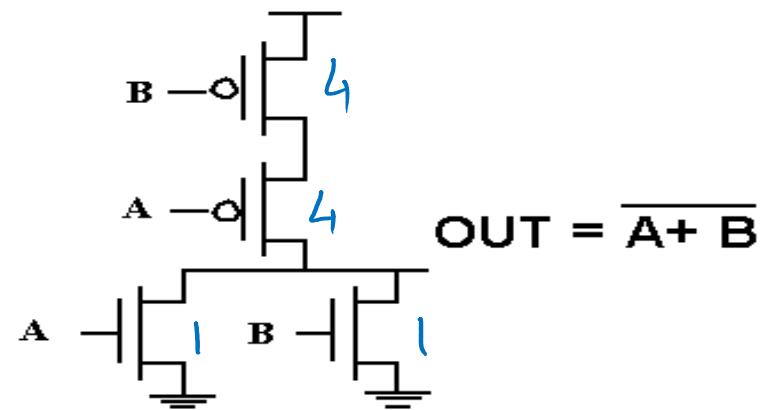
Logical effort is a function of topology, independent of sizing
Effective fanout is a function of load/gate size

Logical Effort of Gates



Delay Of NOR-2 Gate

1. Size for same resistance as inverter
2. LE = ratio of input cap of gate versus inverter



$$LE = \frac{\frac{R_{\text{req, gate}} \cdot C_{\text{in, gate}}}{R_N}}{\frac{R_{\text{req, inv}} \cdot C_{\text{in, inv}}}{R_N}} = \frac{5}{3}$$

Intrinsic capacitance (C_{dnor}) = $6C_d = 6 \cdot g \cdot C_g = 2 \cdot g \cdot C_{\text{in, inv}}$

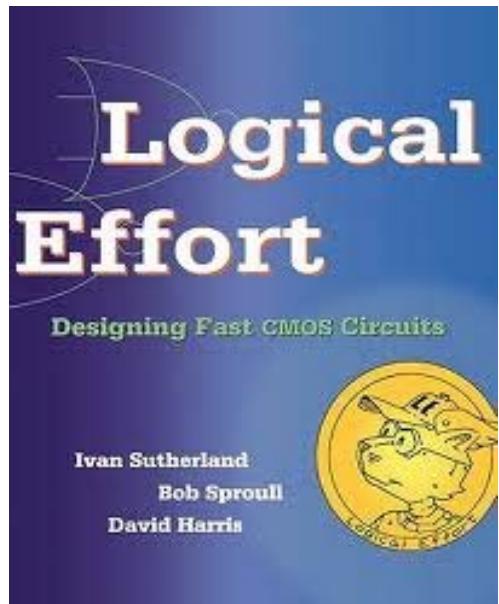
$$t_{\text{pint}}(\text{NOR}) = k R_N \cdot C_{\text{dnor}} = k \cdot R_N \cdot \underbrace{C_{\text{in, inv}} \cdot (2g)}_{P_{\text{NAND,2}}}$$

Question

Any logic function can be implemented using NOR gates only or NAND gates only!

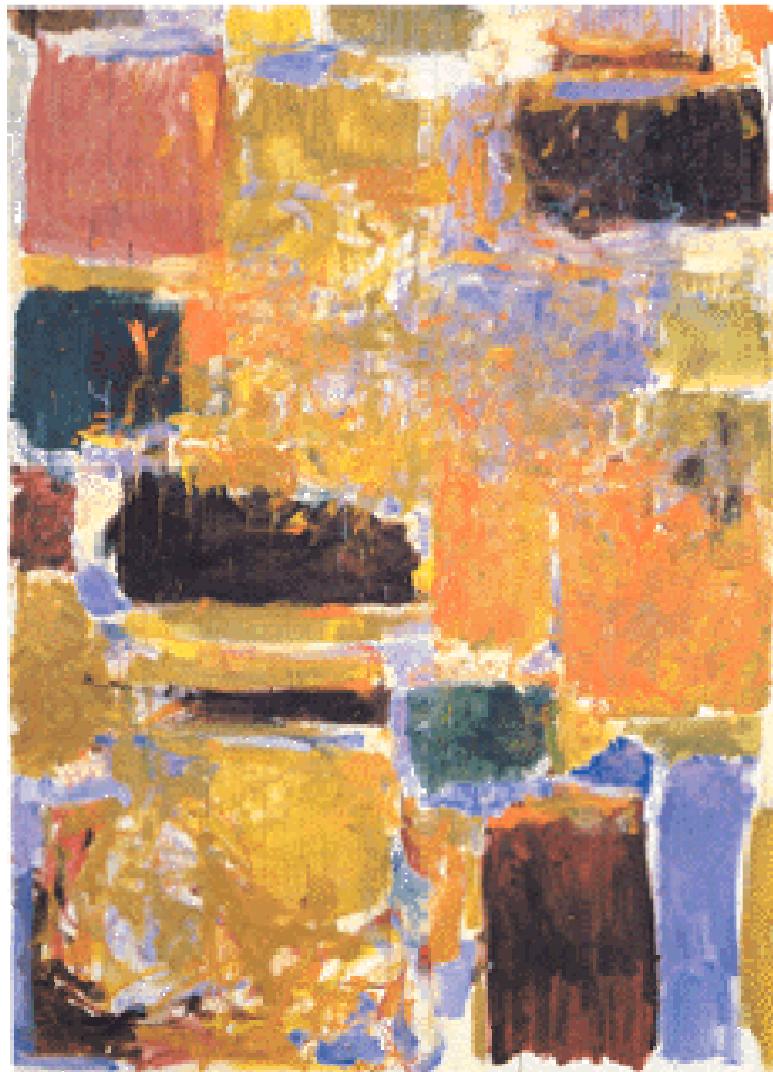
Which of the two approaches is preferable in CMOS (from a performance perspective)?

Logical Effort



Gate Type	Number of Inputs			
	1	2	3	n
Inverter	1			
NAND		4/3	5/3	(n + 2)/3
NOR		5/3	7/3	(2n + 1)/3
Multiplexer		2	2	2
XOR		4	12	

[From Sutherland, Sproull, Harris]



Optimizing Complex Combinational Logic

Multistage Networks

$$Delay = \sum_{i=1}^N (p_i + LE_i \cdot f_i)$$

Effective fanout: $EF_i = LE_i f_i$

Only for tree networks

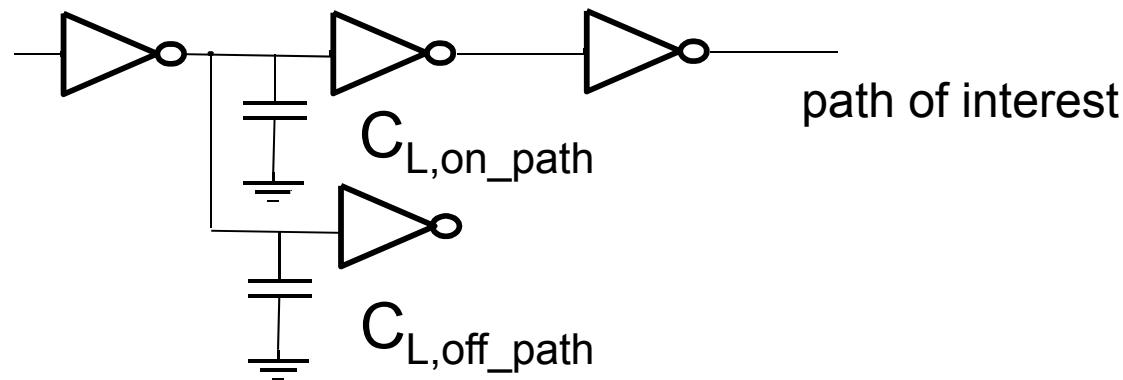
Path delay $D = \sum d_i = \sum p_i + \sum EF_i$

Path electrical fanout: $F = C_L/C_{in} = \prod f_i$

Path logical effort: $\Pi LE = LE_1 LE_2 \dots LE_N$

Path effort: $PE = \Pi LE F$

Adding branching



Branching effort: $b = \frac{C_{L,\text{on-path}} + C_{L,\text{off-path}}}{C_{L,\text{on-path}}}$

Multistage Networks

$$\text{Delay} = \sum_{i=1}^N (p_i + LE_i \cdot f_i)$$

Effective fanout: $EF_i = LE_i f_i$

Path delay $D = \sum d_i = \sum p_i + \sum EF_i$

Path electrical fanout: $F = C_L / C_{in}$

Branching effort: $\Pi B = b_1 b_2 \dots b_N$

$\Pi f_i = \Pi B F$ *(assuming all paths in the tree are important)*

Path logical effort: $\Pi LE = LE_1 LE_2 \dots LE_N$

Path effort: $PE = \Pi LE \Pi B F$

Optimum Effort per Stage

When each stage bears the same effort (effective fanout):

$$E F^N = P E$$

$$E F = \sqrt[N]{P E}$$

Effective fanouts: $LE_1 f_1 = LE_2 f_2 = \dots = LE_N f_N$

Minimum path delay

$$\hat{D} = \sum_{i=1}^N (L E_i f_i + p_i) = N \cdot P E^{1/N} + \sum_{i=1}^N p_i$$

Optimal Number of Stages

For a given load,
and given input capacitance of the first gate
Find optimal number of stages and optimal sizing

$$D = N \cdot P E^{1/N} + \sum p_i$$

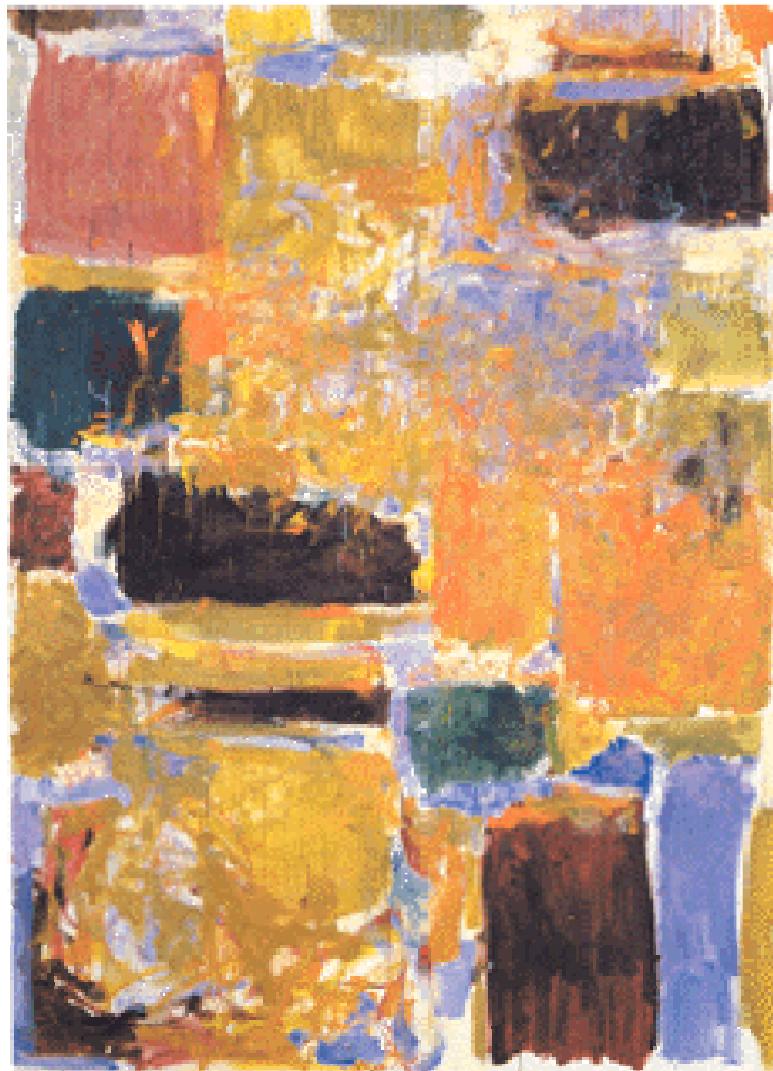
Remember: we can always add inverters to the end of the chain

The ‘best effective fanout’ $EF = P E^{1/\hat{N}}$ is still around 4
(3.6 with $\gamma=1$)

Method of Logical Effort: Summary

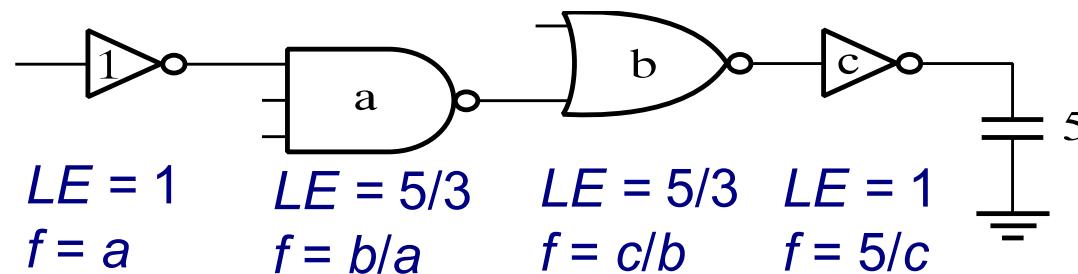
- Compute the path effort: $PE = (\prod LE)BF = EF^{^1}$
- Find the best number of stages $N \sim \log_4 PE = \log_{EF} PE$
- Compute the effective fanout/stage $EF = PE^{1/N}$
- Sketch the path with this number of stages
- Work either from either end, find sizes:
 $C_{in} = C_{out} * LE/EF$

Reference: Sutherland, Sproull, Harris, “Logical Effort”, Morgan-Kaufmann 1999.



Optimizing Complex Combinational Logic: Examples

Example 1: No branching



Electrical fanout, $F =$

$\prod LE =$

PE =

EF/stage =

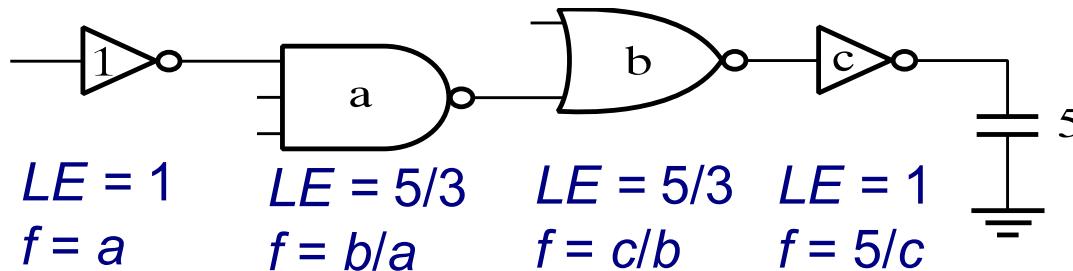
a =

b =

c = $\frac{5 \cdot LE_c}{EF}$

$$\begin{aligned} F &= \\ \frac{5}{3} \cdot \frac{5}{3} &= \frac{25}{9} \\ \frac{125}{9} & \\ PE &= 1/4 \end{aligned}$$

Example 1: No branching



Electrical fanout, $F = 5$

$$\Pi LE = 25/9$$

$$PE = 125/9$$

$$EF/stage = 1.93$$

$$a = 1.93$$

$$b = 2.23$$

$$c = 2.59$$

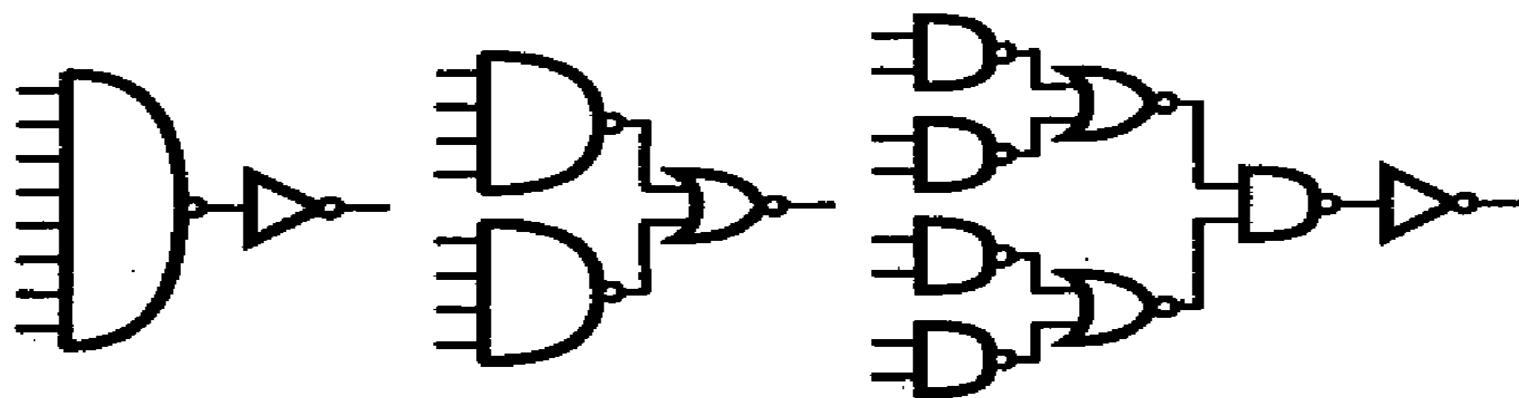
From the back

$$5/c = 1.93$$

$$(5/3)c/b = 1.93$$

$$(5/3)b/a = 1.93$$

Our old problem: which one is better?



$$LE = \frac{10}{3} \quad 1 \\ \Pi LE = \frac{10}{3} \\ P = 8 + 1$$

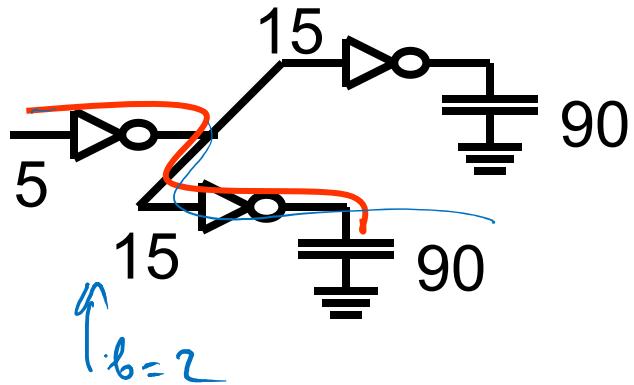
$$LE = 2 \quad \frac{5}{3} \\ \Pi LE = \frac{10}{3} \\ P = 4 + 2$$

$$LE = \frac{4}{3} \quad \frac{5}{3} \quad \frac{4}{3} \quad 1 \\ \Pi LE = \frac{80}{27} \\ P = 2 + 2 + 2 + 1$$

$$\left(\frac{10}{3} \right)^{1/2} \cdot F^{1/2} \quad \left(\Pi LE \cdot F \right)^{1/4} = \left(\frac{80}{27} \right)^{1/4} \cdot F^{1/4} \sqrt{2}$$

43

Adding Branching



$$\begin{array}{rcl} LE & = & 1 \\ F & = & \cancel{90}/\cancel{5} = 18 \\ PE & = & \cancel{18}/6 \text{ (wrong!)} \\ \hline \end{array}$$

$$\begin{array}{rcl} EF_1 & = & (15+15)/5 = 6 \\ EF_2 & = & 90/15 = 6 \\ PE & = & 36, \text{ not } 18! \end{array}$$

Better: $PE = F \cdot LE \cdot B = 18 \cdot 1 \cdot 2 = 36$

Example 2 with Branching

Select gate sizes y and z to minimize delay from A to B

Logical Effort: $LE = \frac{4}{3} \cdot \frac{4}{3} \cdot \frac{4}{3}$

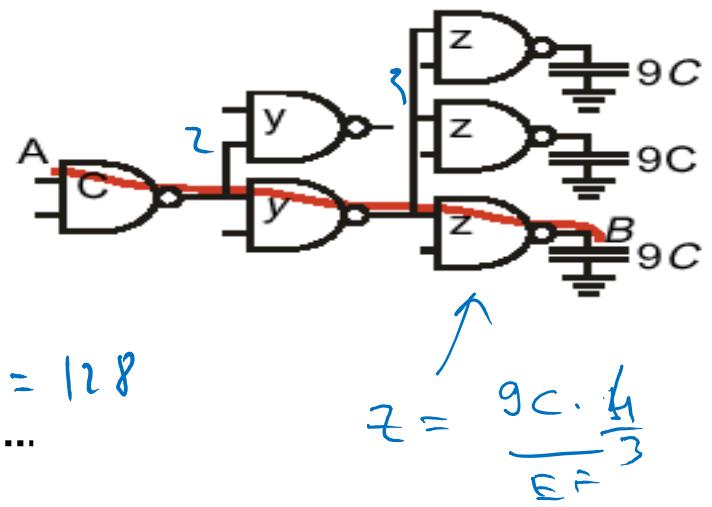
Electrical Fanout: $F = 9$

Branching Effort: $B = 2 \cdot 3$

Path Effort: $PE = 6 \cdot 9 \cdot \left(\frac{4}{3}\right)^3 = 128$

Best Effective Fanout: $EF = \underline{(128)^{1/3}}$ \rightarrow

Delay: $D = \underbrace{P}_{3 \cdot \text{Fanout}-2} + 3 \cdot EF = \underbrace{287}_{\approx 1}$



Example 2 with Branching

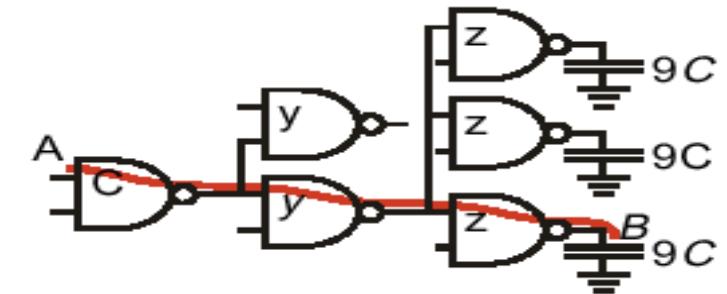
Select gate sizes y and z to minimize delay from A to B

Logical Effort: $LE = (4/3)^3$

Electrical Fanout: $F = C_{\text{out}}/C_{\text{in}} = 9$

Branching Effort: $B = 2 \cdot 3 = 6$

Path Effort: $PE = \prod LE \cdot FO \cdot B = 128$



Best Effective Fanout: $EF = PE^{1/3} \approx 5$

Delay: $D = 3 \cdot 5 + 3 \cdot 2 = 21$

Work backward for sizes:

$$z = \frac{9C \cdot (4/3)}{5} = 2.4C$$

$$y = \frac{3z \cdot (4/3)}{5} = 1.9C$$