EE141-Spring 2012 Digital Integrated Circuits

Lecture 5
Switch Logic - Optimization

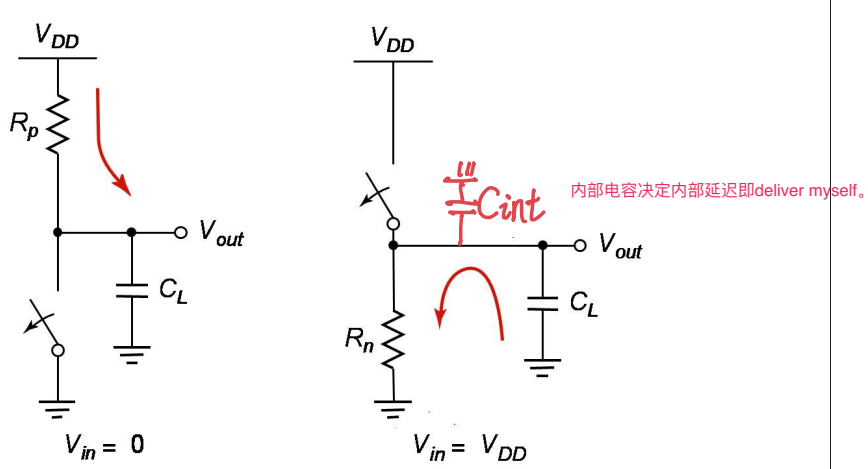
Administrativia

- Homework #2 posted
- First lab done by now – Lab 2 next week

Class Material

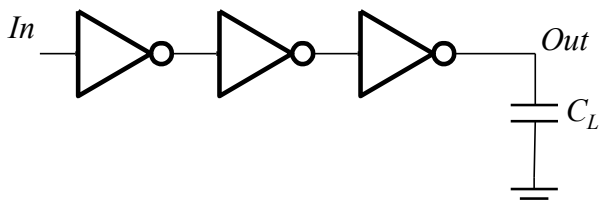
- Last lecture
  - Introduction to switch logic
- Today's lecture
  - Logic optimization
- Reading (5.4-5.5)

The Inverter Switch Model



(a) Low-to-high (b) High-to-low

Driving Capacitance



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

Inverter Delay Model

Delay Formula

$$Delay \sim R_W \left(C_{int} + C_L \right)$$

Cd

designer可以影响的因子仅有一个，那就是f

$$t_p = kR_W C_{\text{in}} \left(C_{\text{int}} / C_{in} + C_L / C_{in} \right) = t_{inv} \left(\gamma + f \right)$$

Cg

$C_{int} = \gamma\, C_{in}$ ($\gamma \approx 1$ for CMOS inverter) 不是一个常数但一般约等于1

$f = C_L / C_{in}$ – electrical fanout

$t_{\text{inv}} = kR_{min} C_{ginv}$ = the delay of minimum size inverter without any CL=inverter本身的延迟

任何inverter的内部延迟都是一个常数，仅与technology有关。

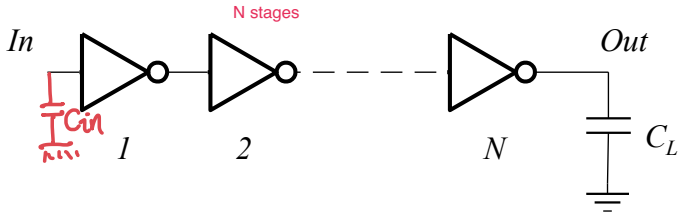
t_{inv} is independent of sizing of the gate!!!

Inverter Optimization Problem #1

Delay Optimization Problem #1

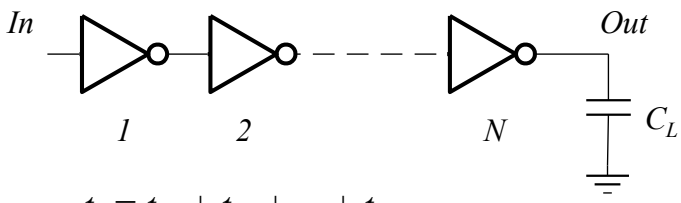
- You are given:
  - A fixed number of inverters
  - The size of the first inverter
  - The size of the load that needs to be driven
- Your goal:
  - Minimize the delay of the inverter chain
- Need model for inverter delay vs. size

Apply to Inverter Chain



在这里重新定义公式中的f，公式中的f用反相器的自身的负载电容/反相器自身的栅极电容代替；而f用于表示一个门的fanout。F= C_L/C_{in} 这两个电容分别是第一个反相器的栅极电容和最后一个反相器的负载电容。F是整个inverter chain的total fanout。

Apply to Inverter Chain



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

$$t_{pj} = t_{inv}\left(\gamma + \frac{C_{in,j+1}}{C_{in,j}}\right) \begin{matrix} = C_L \\ = C_{in} \end{matrix}$$

上下两个电容分别为第j个反相器的栅极电容和负载电容

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

Optimal Sizing for Given N

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

偏微分

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

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

Optimal Sizing for Given N (cont'd)

- Result: every stage has **equal fanout**:

$$f_{j-1} = \boxed{\frac{C_{in,j}}{C_{in,j-1}} = \frac{C_{in,j+1}}{C_{in,j}}} = f_j \text{ 考虑 } \frac{C_L}{C_{in,N-1}} \Rightarrow f^N = F$$

- In other words, size of each stage is geometric mean of two neighbors:

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

- Equal fanout → every stage will have same delay

对于同一条inverter chain，F应该拥有同一个数值，它是一个常数。

Optimum Delay and Number of Stages

why?因为CL和Cin是确定的，我们不能改变它们，但是又要确保每个stage的delay相同，确保每个stage都有相同的工作量，同时确保最后一个反相器要足够大去驱动CL，所以需要divide and conquer，使每个inverter以f倍递增，不能使每个inverter都相同，即 $f^N = F$

- When each stage has same fanout f:

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

因此第一个设为1，第二个设为f，第三个设为f的平方，。。。，第N个设为f的n-1次方

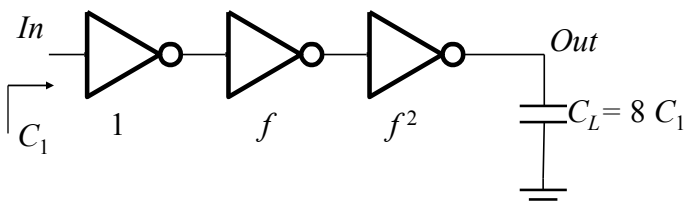
- Effective fanout of each stage:

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

- 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:

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

Inverter Optimization Problem #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:

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

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

Untangling the Optimization Problem

$f = \left(\frac{C_L}{C_{in}}\right)^{\frac{1}{N}}$

Untangling the Optimization Problem

- Rewrite N in terms of fanout/stage f:

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

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

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

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

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

这里是Cd=0的情况，Cg永远不可能为0，因为它是分母。但这种情况其实永远不会发生。

没有直接给出最优的stages数量（N），但可以通过计算得到。
不过这个方程直接计算也比较困难，需要通过Matlab计算或者画图。

The Exponential Horn

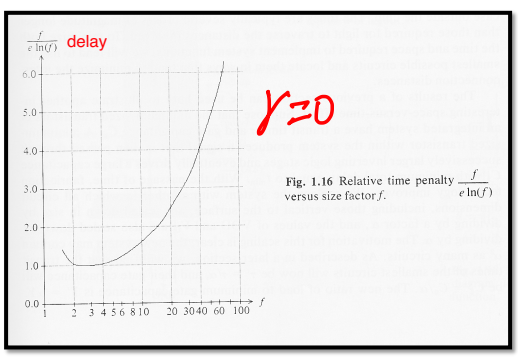


Fig. 1.16 Relative time penalty $\frac{f}{e\ln(f)}$ versus size factor f.

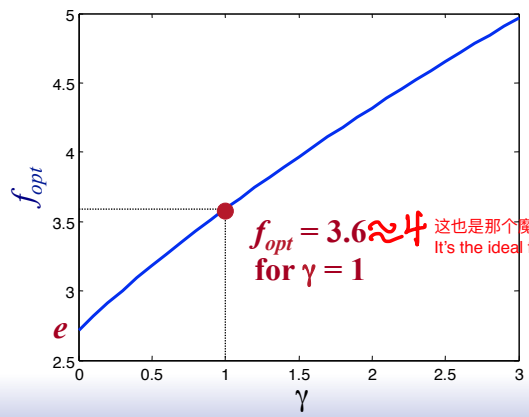
$\gamma = 1$
$f = E$

Mead-Conway, 1980

Optimum Effective Fanout f

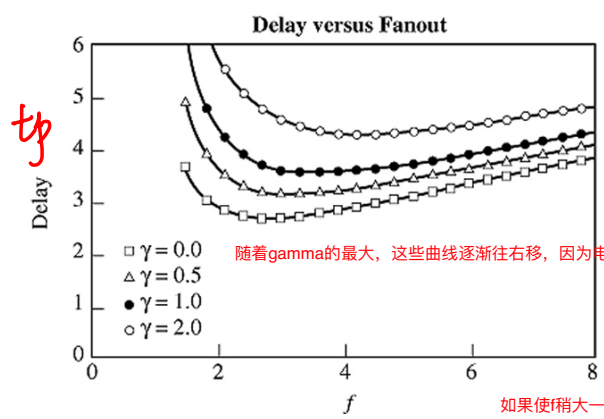
- Optimum f for given process defined by γ

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



这也是那个魔幻的数字4的由来。
It's the ideal fanout factor for basically driving.

In Practice: Plot of Total Delay



[Hodges, p.281]

随着gamma的最大，这些曲线逐渐往右移，因为电容变大了。

- Curves very flat for $f > 2$

如果使f稍大一点，在layer上的影响会非常小
因为在上图曲线的右边，曲线比较平坦，但
是左边比较陡峭。

  - Simplest/most common choice: $f = 4$

Normalized Delay As a Function of F

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

one stage

F	Unbuffered N=1	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

delay (单位;tinv)

反相器数量越多，所需要的能量越多，所需要的面积也越大，因为这些反相器数量变多了，而且逐级增大，但在一定范围内（F的不同区间和N的不同区间），反相器数量越多，delay越小。

$(\gamma = 1)$

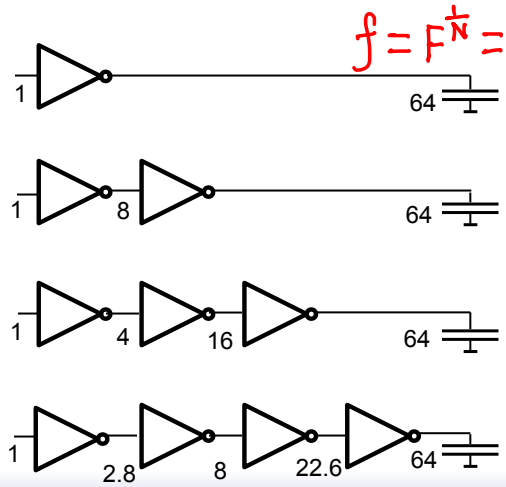
Textbook: page 210

ebook:189

$E = C_L V_{DD}^2$

Buffer Design

$f = F^{\frac{1}{N}} = 64^{\frac{1}{N}}$



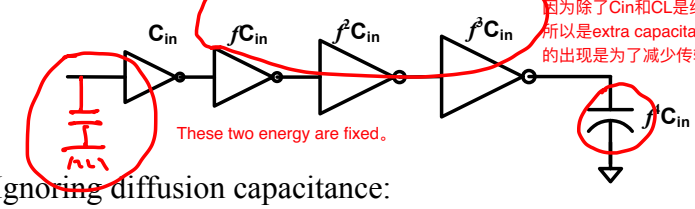
N	f	t_p
1	64	65
2	8	18
3	4	15
4	2.8	15.3

delay变大因为overhead变得太大。

What About Energy (and Area)?

What About Energy (and Area)?

extra capacitance（注：也对应额外能量）。因为除了Cin和CL是给定的，其它都是未知的，所以是extra capacitance。简单点讲，额外电容的出现是为了减少传输延迟，也是为了驱动CL。



Ignoring diffusion capacitance:

$$C_{tot} = C_{in} + f \cdot C_{in} + \ldots + f^N \cdot C_{in}$$
$$= C_{in} \cdot (1 + f + \ldots + f^N)$$
$$= C_{in} + C_{in} \cdot f^N + C_{in} \cdot \underbrace{f \cdot (1 + f + \ldots + f^{N-2})}_{\text{Overhead !!! } f(f^{N-1}-1)/(f-1)}$$

等比数列求和

C_L

Example (γ=0): C_L = 20pF; C_i = 50fF → N = 6

Fixed: 20pF

Overhead: 11.66pF !!!

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

m:-3 u:-6 n:-9 p:-12 f:-15

如果延迟允许变大一点，我们可以通过把last stage的电容适当缩小一点从而减少面积和能量的使用。

要改变某些东西，往往要改变的是占据主导地位或者控制地位的因素，其它小因素往往不需要关注。

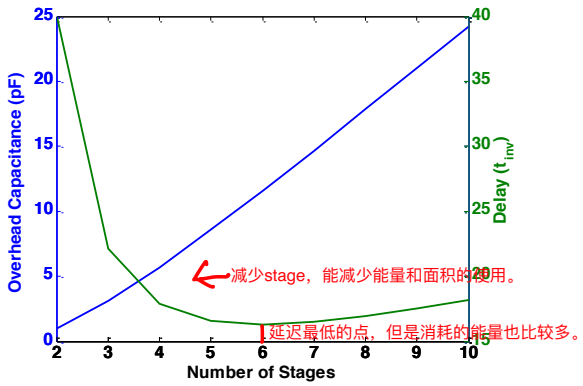
$f = e^{(1+\gamma/f)}$

$\gamma = 0$

$\gamma = \frac{C_{int}}{C_{in}}$

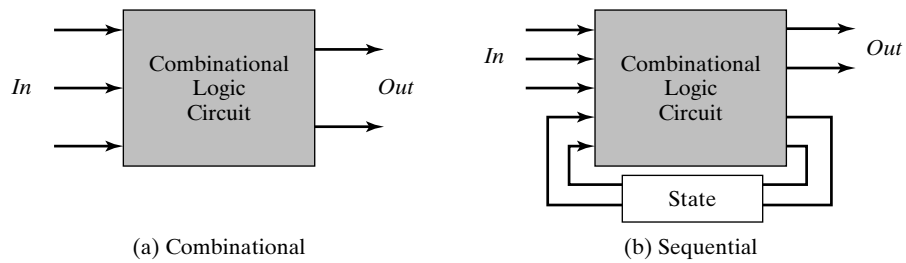
Example Overhead Numbers

Example: C_L = 20pF; C_{in} = 50fF



Moving to Logic

Combinational vs. Sequential Logic



(a) Combinational (b) Sequential

Output = *f*(*In*) **Output = *f*(*In, Previous In*)**

Static Logic Gates

At every point in time (except during the switching transients) each gate output is connected to either V_{DD} or V_{SS} via a low resistive path. 指门停留在稳态

ground

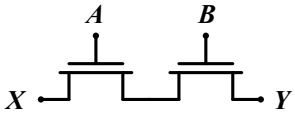
The outputs of the gates assume at all times the value of the Boolean function implemented by the circuit (ignoring, once again, the transient effects during switching periods).

(Will contrast this later to dynamic circuit style.)

Switch Logic

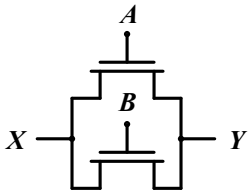
- Transistor ↔ switch controlled by its gate signal
  - NMOS switch closes when switch control input is high

AND



$Y = X$ if A AND B

OR

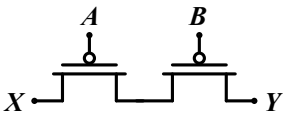


$Y = X$ if A OR B

Switch Logic - Continued

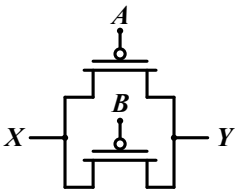
- PMOS switch closes when switch control is low

NOR



$Y = X$ if $\overline{A}$ AND $\overline{B} = \overline{A + B}$

NAND



$Y = X$ if $\overline{A}$ OR $\overline{B} = \overline{AB}$

单一MOS管构造的门或非对偶构造的门的特点：
只对输出电压为1时有定义，对输出电压为0无定义，当MOS管导通时（稳态），
输入端和输出端/地之间存在电阻，一直有电流消耗即静态功耗。

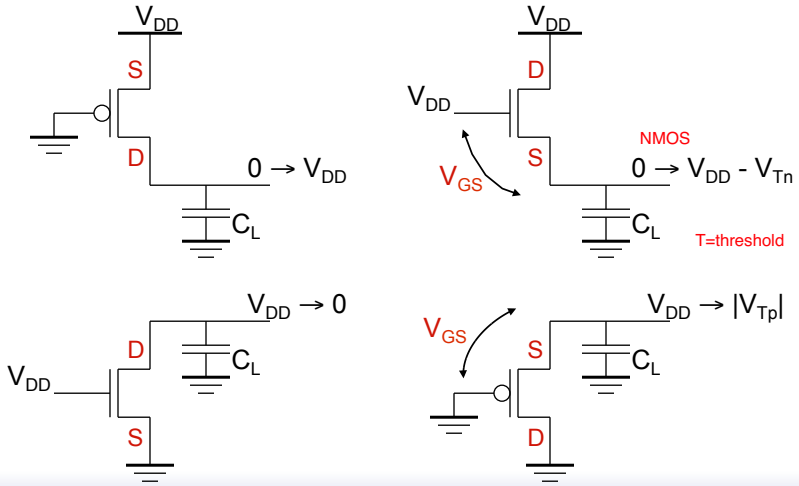
为什么不是NMOS在上面，PMOS在下面？因为这样的话，对偶会丢失。

Some Limitations

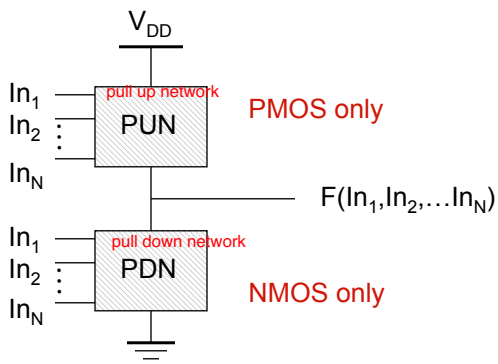
When to use NMOS or PMOS transistors?

Limitations

NMOS输出低电压很理想，但输出高电压不理想
PMOS输出高电压很理想，但输出低电压不理想；
因此才需要对偶。



Static Complementary CMOS



PUN and PDN are dual logic networks one is on another is off。
PUN and PDN functions are complementary

Complementary CMOS Logic Style

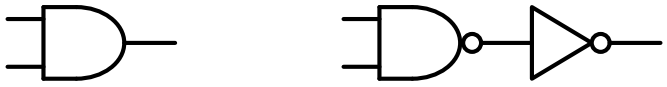
- PUP is the **dual** to PDN (parallel)
  (can be shown using DeMorgan's Theorems)

$$\overline{A+B} = \overline{A}\,\overline{B}$$

$$\overline{AB} = \overline{A} + \overline{B}$$

- Static CMOS gates are always inverting

CMOS logic is always inverting by nature.



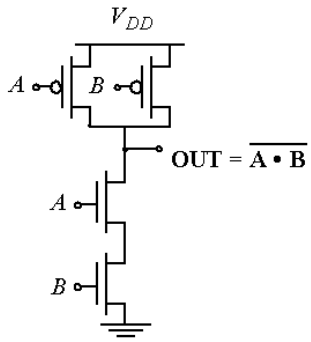
AND = NAND + INV

PMOS实现逻辑门的1，NOMS实现逻辑门的0.

Example Gate: NAND

A	B	Out
0	0	1
0	1	1
1	0	1
1	1	0

Truth Table of a 2 input NAND gate

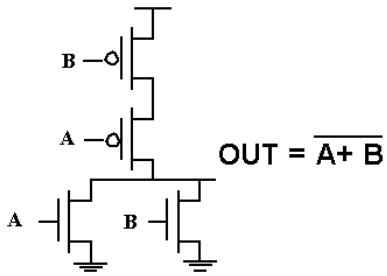


- PDN: G = AB ⇒ Conduction to GND
- PUN: F = $\overline{A}$ + $\overline{B}$ = $\overline{AB}$ ⇒ Conduction to V_{DD}
- $\overline{G(In_1, In_2, In_3, \ldots)} \equiv F(\overline{In_1}, \overline{In_2}, \overline{In_3}, \ldots)$

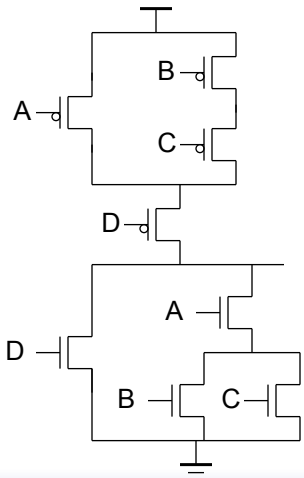
Example Gate: NOR

A	B	Out
0	0	1
0	1	0
1	0	0
1	1	0

Truth Table of a 2 input NOR gate



Complex CMOS Gate



$$OUT = \overline{D + A \bullet (B + C)}$$

CMOS Properties

- Full rail-to-rail swing
- Symmetrical VTC
- Propagation delay function of load capacitance and resistance of transistors
- No static power dissipation
- Direct path current during switching