

计算机体系结构

胡伟武、汪文祥

二进制与逻辑电路

- 计算机中数的表示
- CMOS门电路及工艺
- CMOS组合逻辑与时序逻辑
- CMOS电路延迟
- 从Verilog到GDSII
- 其它“0”和“1”表示方法

计算机中数的表示

计算机中数的表示

- 二进制
 - 最容易逻辑实现
 - 自然界中的二值系统较多
- “1”和“0”的表示
 - 用电压的高低表示，半导体工艺，CMOS
 - 用磁通量的有无表示，超导体工艺
 - 用能级的高低表示，量子计算机
 - 用基因序列表示，A, G, C, T, DNA计算机，非二进制？

定点数的表示 (1)

- 原码: $A = a_{n-1} a_{n-2} \dots a_1 a_0$ 表示
 - 最高位 a_{n-1} 为符号位, 0 表示正, 1 表示负。
 - 其它位 $a_{n-2} \dots a_1 a_0$ 表示数值。
 - 原码的问题: 加减法效率低, 两个 “0”
- 补码
 - 本质是取模运算, 如 $-2 \% 12 = 10$
 - 最高位 a_{n-1} 为符号位, 0 表示正, 1 表示负。
 - $A = a_{n-1} a_{n-2} \dots a_1 a_0$ 表示 $(-2^{n-1} a_{n-1} + a_{n-2} \dots a_1 a_0)$
 - $a_{n-1} = 0$ 时, 补码和原码一样, A 表示正 $a_{n-2} \dots a_1 a_0$ 。
 - $a_{n-1} = 1$ 时, A 表示 $(-2^{n-1} + a_{n-2} \dots a_1 a_0)$ 。
- 原码与补码的转换
 - 最高位为 0 时, 一样
 - 最高位为 1 时, 最高位不变, 其余位 “按位取反加一”。

定点数的表示（2）

- 补码运算

- 取负数，每一位(包括符号位)求补，即按位取反加一。
- $A-B=A+B$ 的负数= $A+(B$ 求补)
- 加法溢出判断：A和B的最高位一样，且结果的最高位与A和B的最高位不一样。
- $1001+0101(-7+5)=1110$, $1100+0100(-4+4)=0000$,
- $0011+0100(3+4)=0111$, $1100+1111(-4-1)=1011$,
- $0101+0100(5+4)=1001$, $1001+1010(-7-6)=0011$

	+18	-18
8 位 原 码	0 0 0 1 0 0 1 0	1 0 0 1 0 0 1 0
8 位 补 码	0 0 0 1 0 0 1 0	1 1 1 0 1 1 1 0
16 位 原 码	0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0	1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0
16 位 补 码	0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0	1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 0

浮点数的表示 (1)

- 定点数的不足
 - 表示范围有限，太大或太小的数都不能表示
 - 除法不精确
- 浮点数的表示：IEEE 754标准
 - 三部分组成：符号位，阶码(exponent)，尾数(fraction)
 - 最高位是符号位
 - 阶码的移码表示，底为2， $2^{(\text{阶码}-\text{偏移值})}$
 - 规格化表示，尾数的最高位总为1，因此可以不存
 - 单精度和双精度
 - 扩展的单双精度

S	Exp. (8)	Fraction(23)
---	----------	--------------

S	Exp. (11)	Fraction(52)
---	-----------	--------------

浮点数的表示 (2)

- IEEE 754 浮点格式参数

参 数	单 精 度	扩 展 单 精 度	双 精 度	扩 展 双 精 度
字 长	32	≥ 43	64	≥ 79
阶 码 位 数	8	≥ 11	11	≥ 15
阶 码 偏 移 量	127	未 定	1023	未 定
最 大 阶 码	127	≥ 1023	1023	≥ 16383
最 小 阶 码	-126	≤ -1022	-1022	≤ -16382
表 示 范 围	$10^{-38}, 10^{+38}$	未 定	$10^{-308}, 10^{+308}$	未 定
尾 数 位 数	23	≥ 31	52	≥ 63
阶 码 个 数	254	未 定	2046	未 定
尾 数 个 数	2^{23}	未 定	2^{52}	未 定
值 的 个 数	$1.98 * 2^{31}$	未 定	$1.99 * 2^{63}$	未 定

浮点数的表示 (3)

- IEEE 754 浮点格式

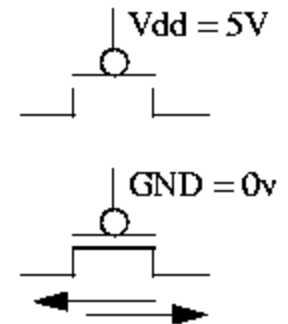
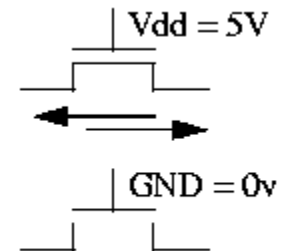
	单精度				双精度			
	符号位	阶码	尾数	值	符号位	阶码	尾数	值
正无限	0	255	0	∞	0	2047	0	∞
负无限	1	255	0	$-\infty$	1	2047	0	$-\infty$
Quiet NaN	0 或 1	255	高位=0	NaN	0 或 1	2047	高位=0	NaN
Signaling NaN	0 或 1	255	高位=1	NaN	0 或 1	2047	高位=1	NaN
正规格化非 0	0	$0 < e < 255$	f	$2^{e-127}(1.f)$	0	$0 < e < 2047$	f	$2^{e-1023}(1.f)$
负规格化非 0	1	$0 < e < 255$	f	$-2^{e-127}(1.f)$	1	$0 < e < 2047$	f	$-2^{e-1023}(1.f)$
正非规格化非 0	0	0	$f \neq 0$	$2^{e-126}(0.f)$	0	0	$f \neq 0$	$2^{e-1022}(0.f)$
负非规格化非 0	1	0	$f \neq 0$	$-2^{e-126}(0.f)$	1	0	$f \neq 0$	$-2^{e-1022}(0.f)$
正 0	0	0	0	0	0	0	0	0
负 0	1	0	0	-0	1	0	0	-0

- 把 Singaling NaN 当作操作数引起例外, 把 Quiet NaN 当作操作数结果也为 Quiet NaN, 可以用尾数部分区分它们。
- 非规格化数用于阶码下溢的情况, 填补最小数和 0 之间的“空隙”。

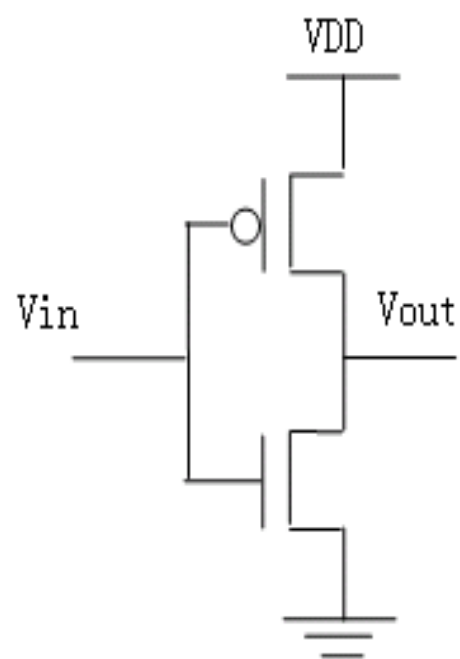
CMOS电路及工艺

CMOS电路---P管与N管

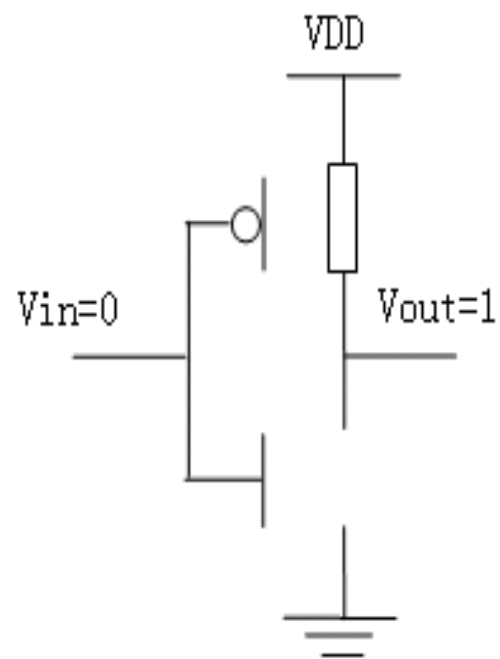
- **CMOS: Complementary Metal Oxide Semiconductor**
 - **NMOS: N-type Metal Oxide Semiconductor**
 - **PMOS: P-type Metal Oxide Semiconductor**
- **NMOS管**
 - 门电压为高时导通
 - 门电压为低时关闭
- **PMOS管**
 - 门电压为高时关闭
 - 门电压为低时导通



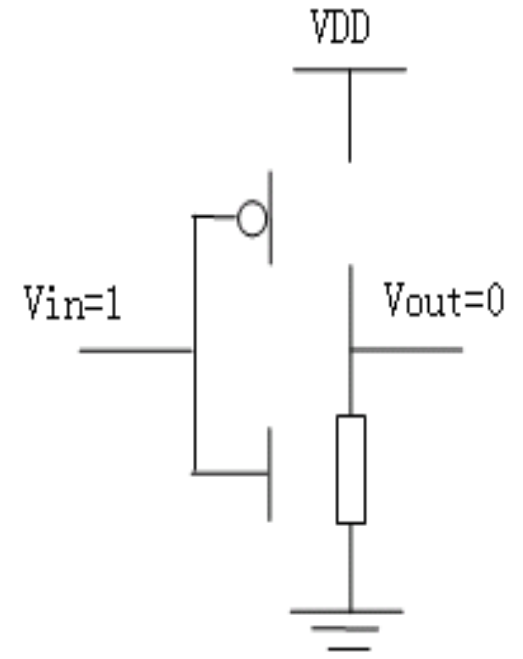
CMOS电路----反相器



(a)



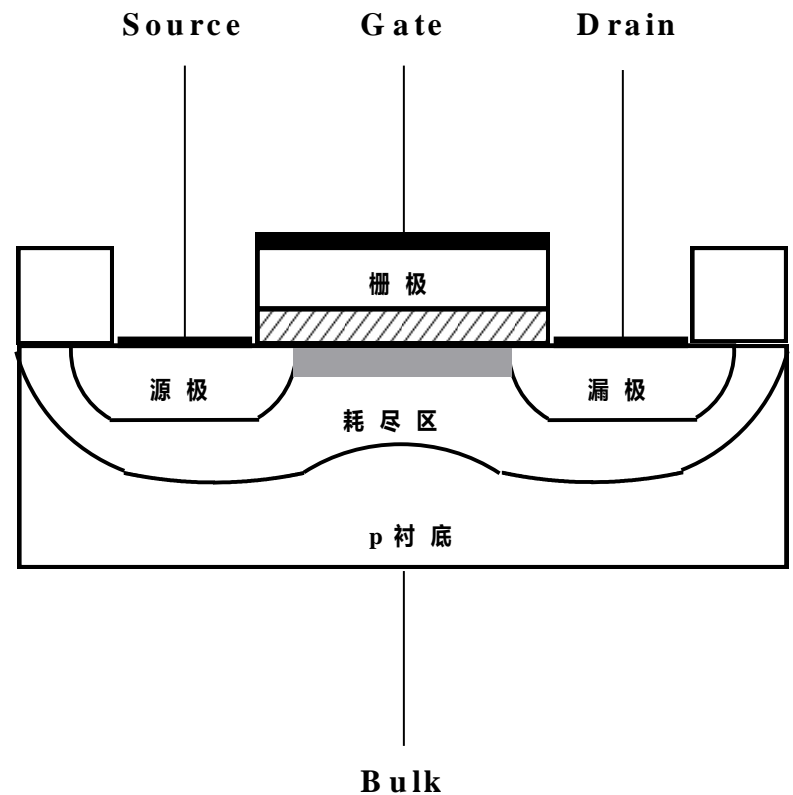
(b)



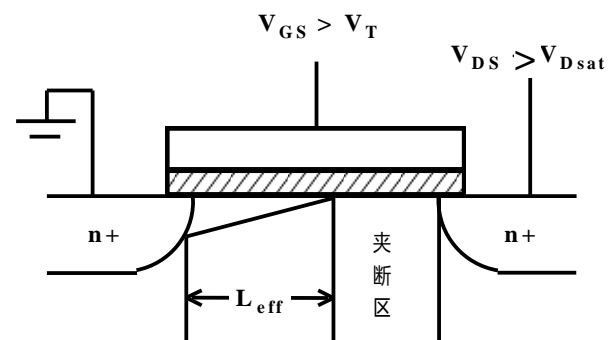
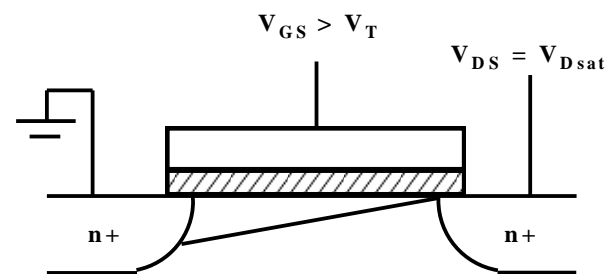
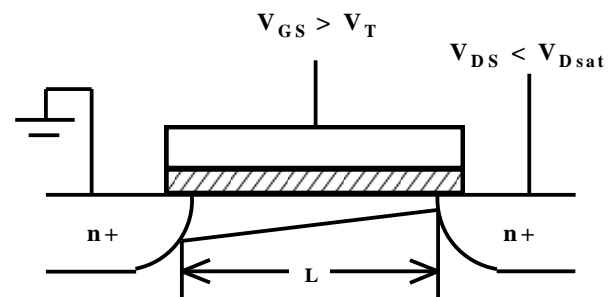
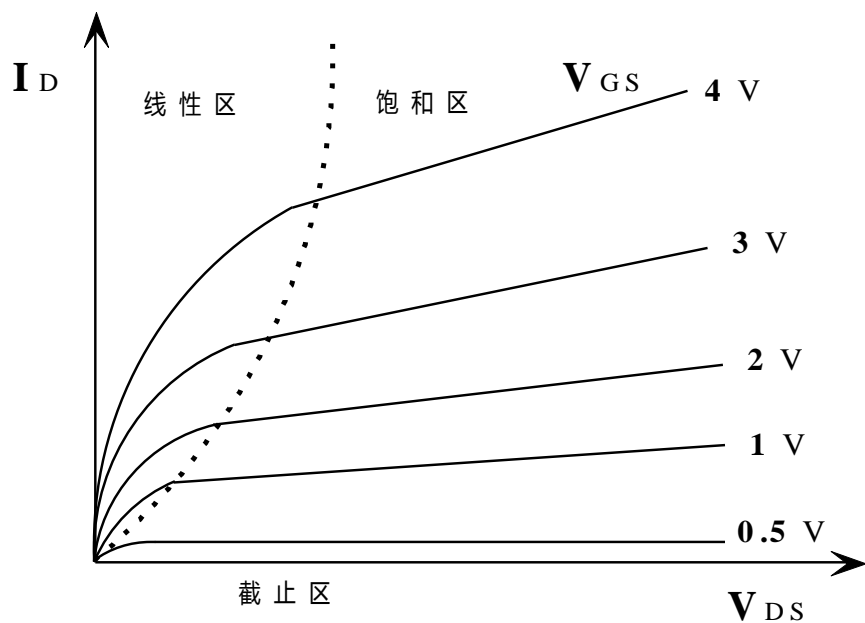
(c)

N沟道MOS晶体管的示意图

- 栅极不加电时，源漏之间是一对PN结，不导通
- 栅极加电后，吸引电子到栅氧化层下面，形成导电的沟道。
 - 附近形成耗尽区



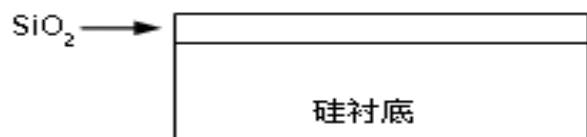
MOS晶体管的工作状态



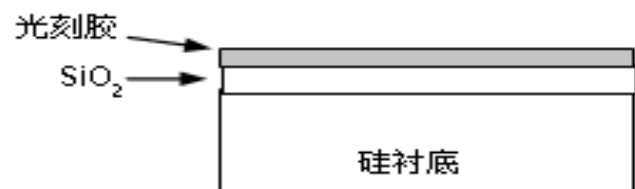
MOS基本工艺---光刻



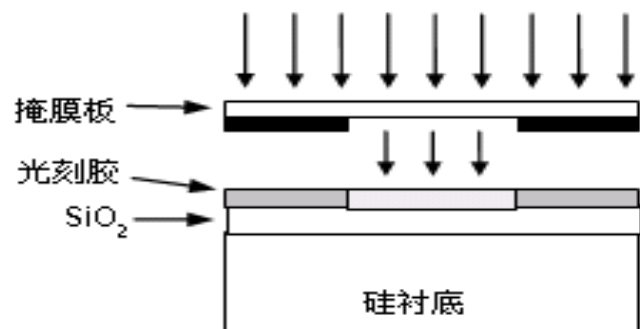
(a)



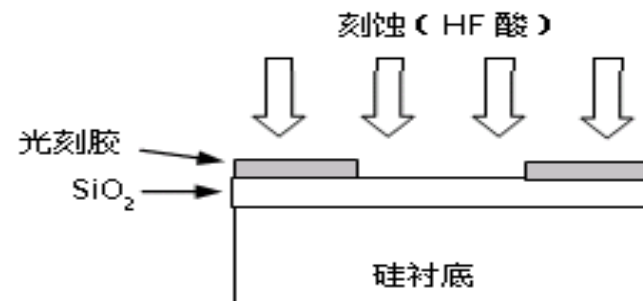
(b)



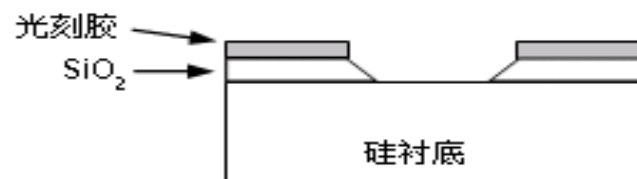
(c)



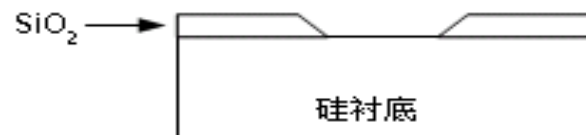
(d)



(e)

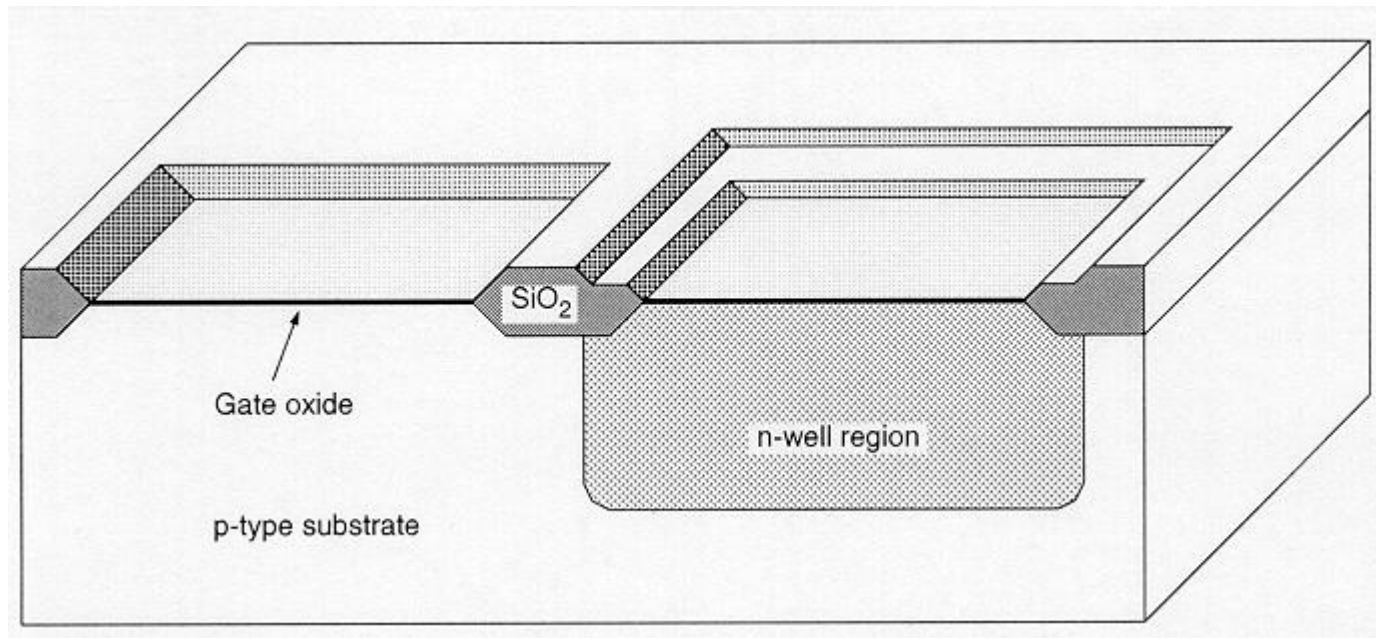


(f)

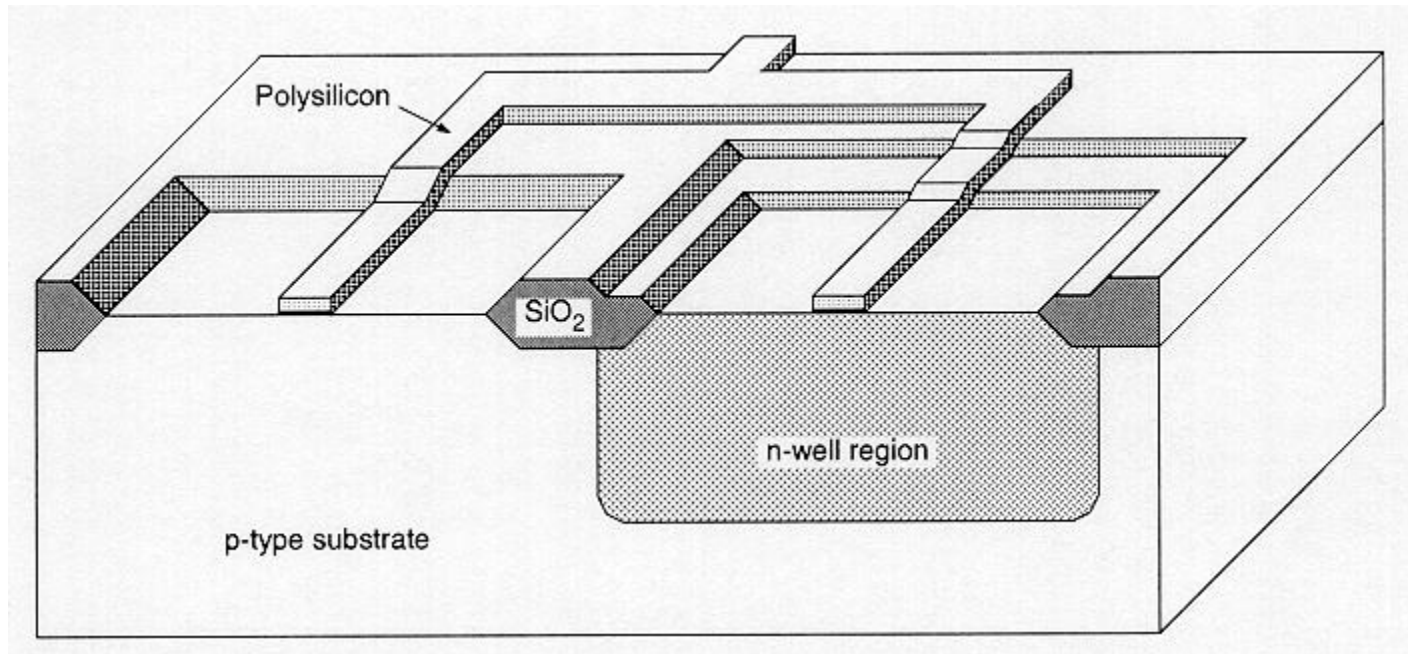


(g)

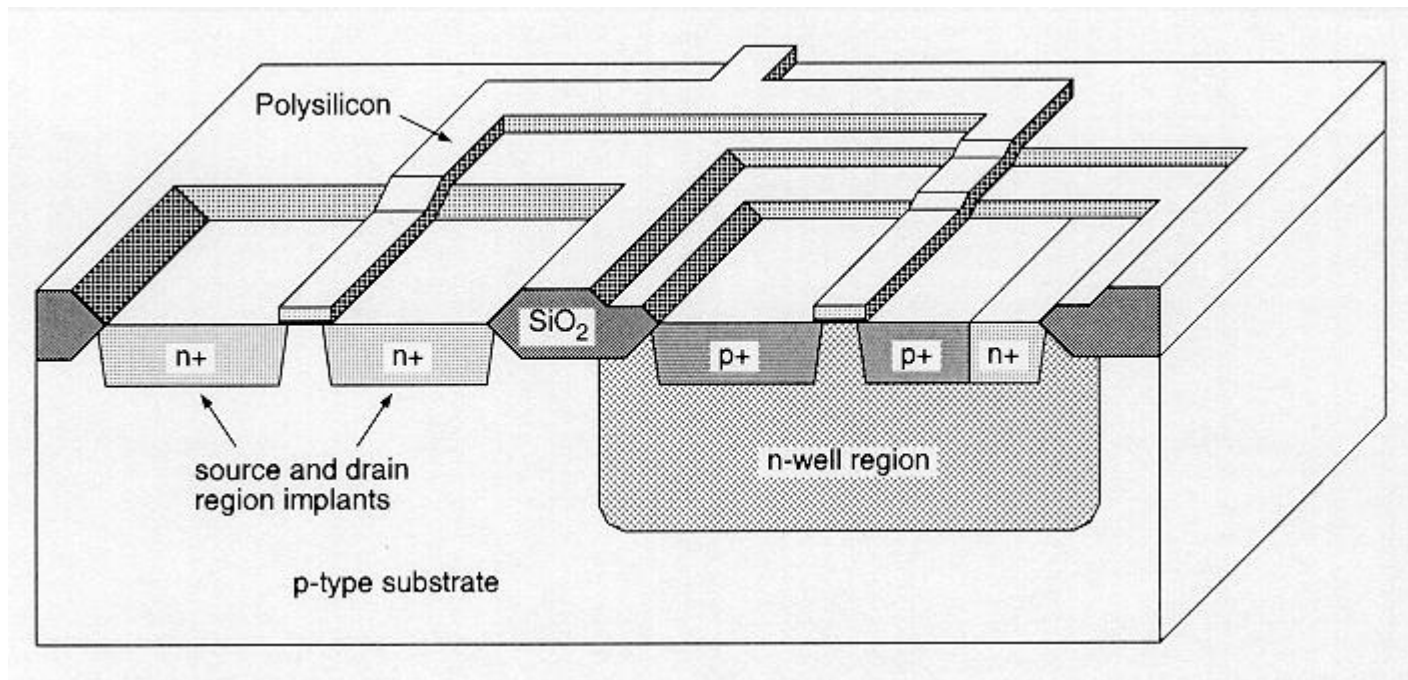
P衬底 nWell CMOS工艺 (1)



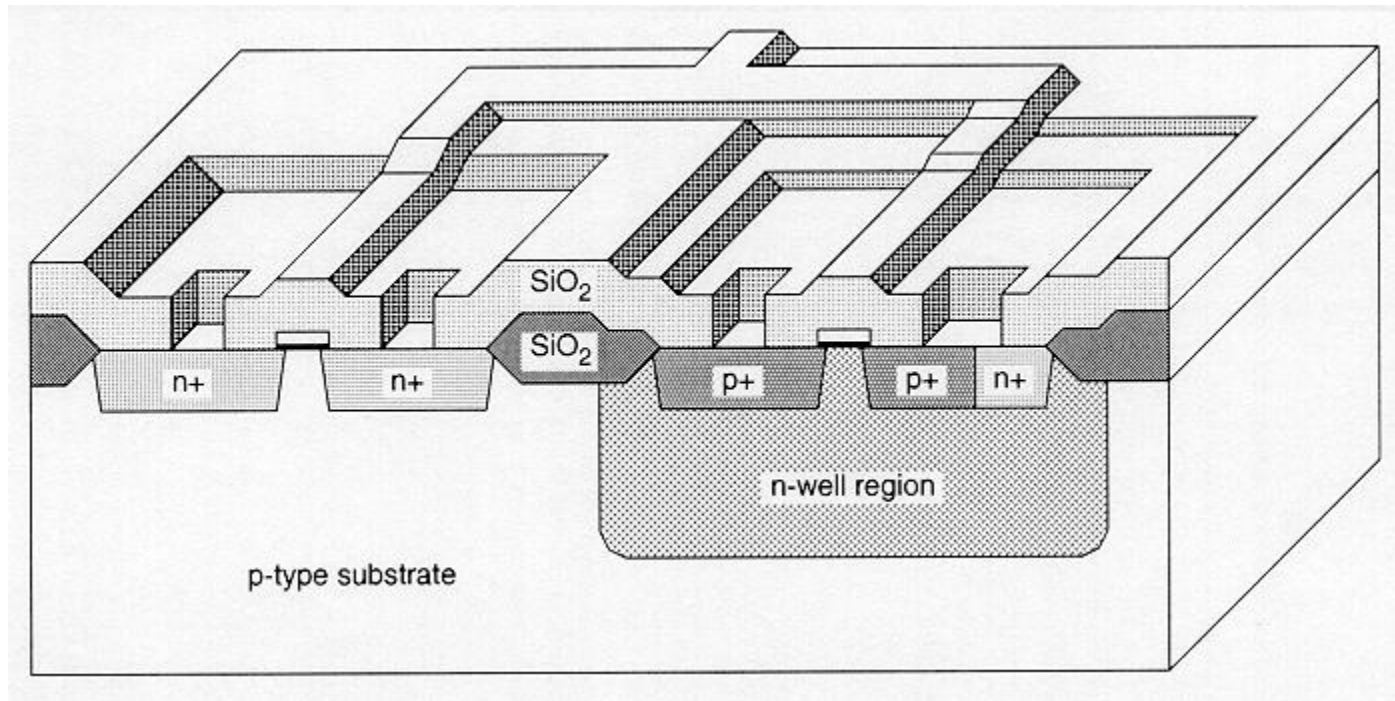
P衬底 nWell CMOS工艺 (2)



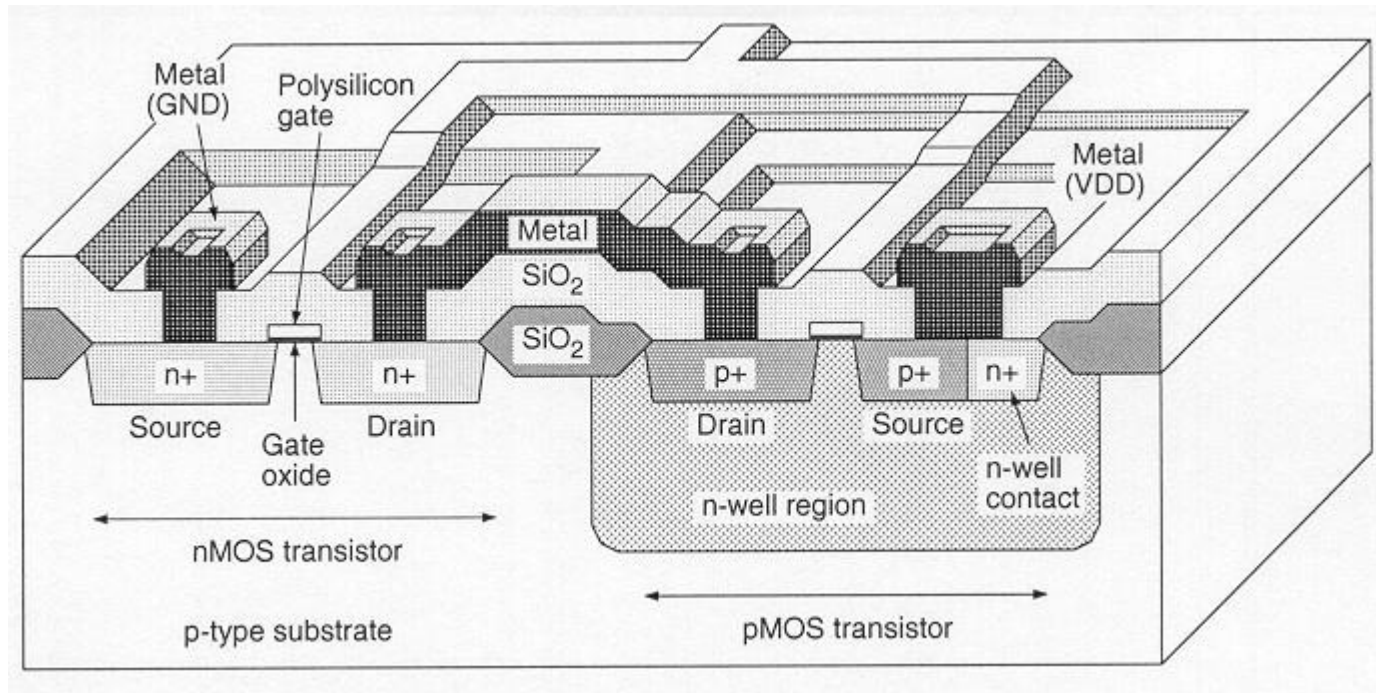
P衬底 nWell CMOS工艺 (3)



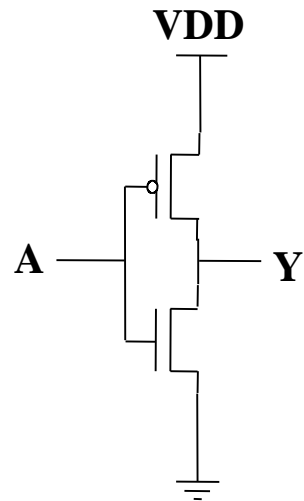
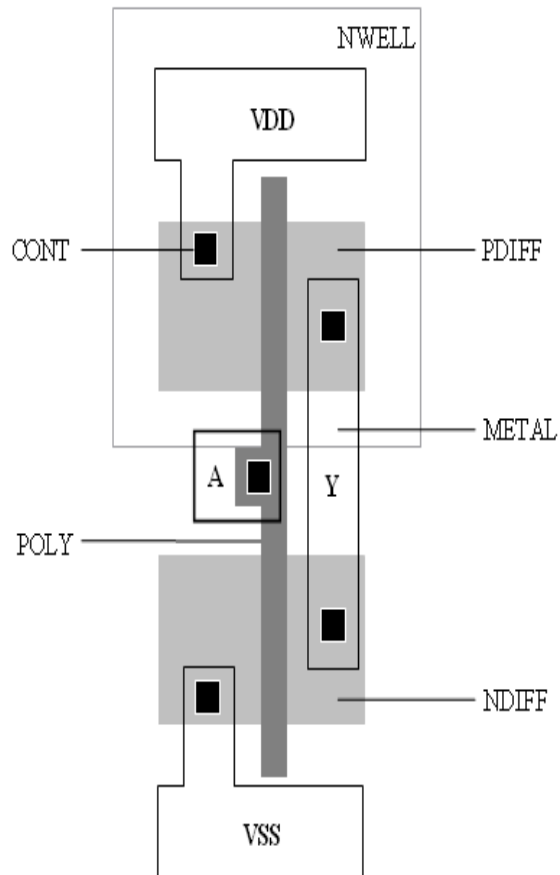
P衬底 nWell CMOS工艺 (4)



P衬底 nWell CMOS工艺 (5)

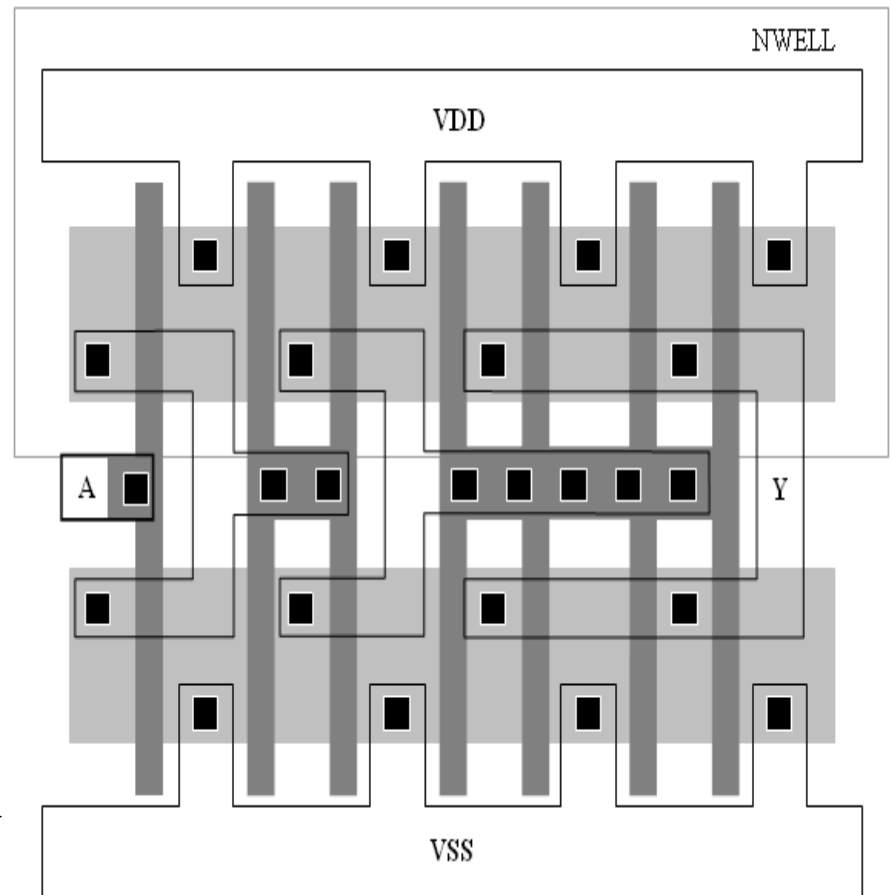


反相器的版图

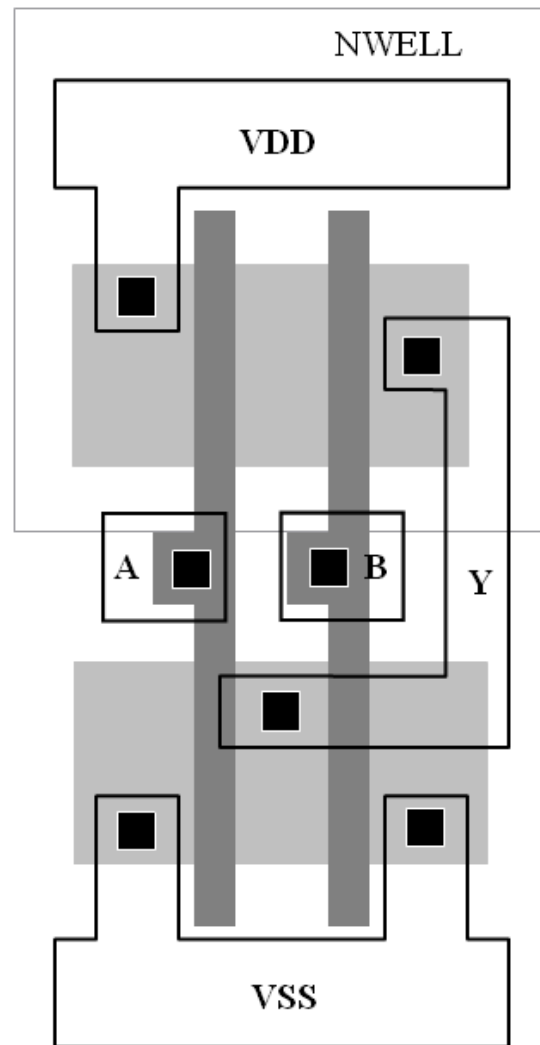
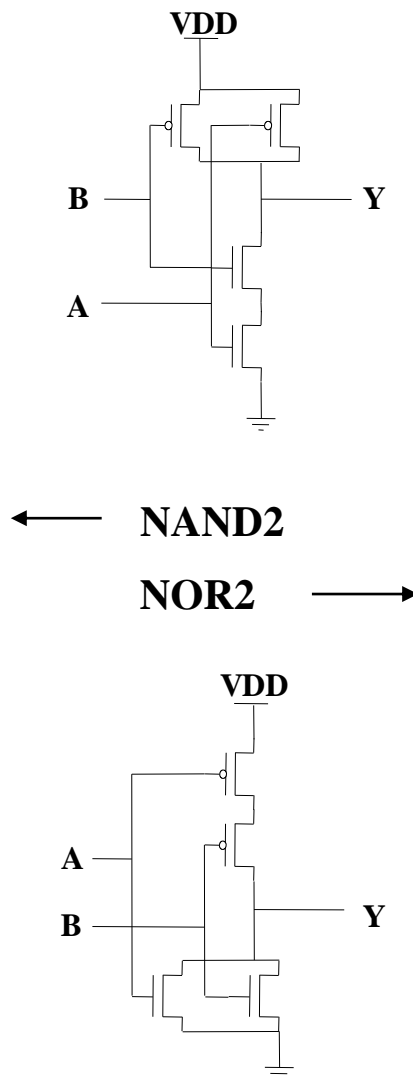
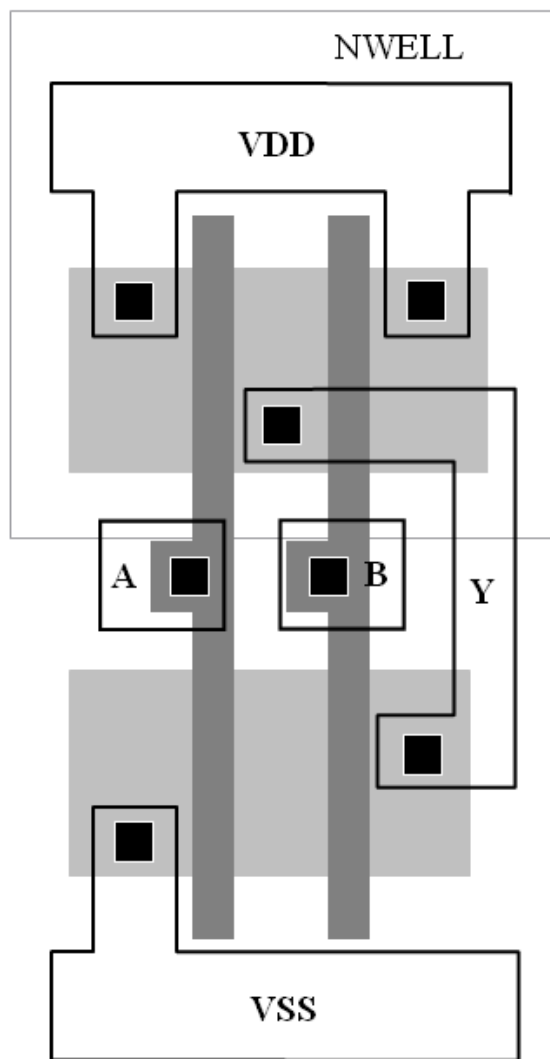


← INVX1

INVX8 →

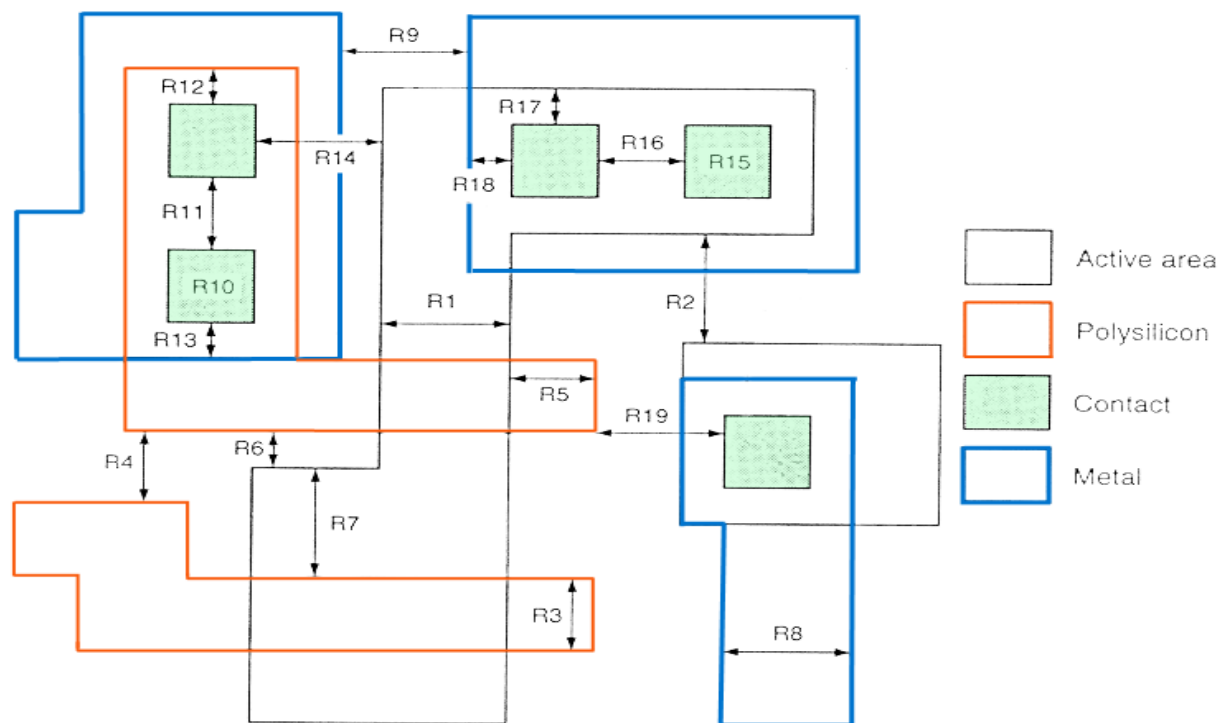


NAND2 和 NOR2 的电路及版图



版图的设计规则

R1	Minimum active area width	0.22
R2	Minimum active area spacing	0.38
R3	Minimum poly width	0.18
R4	Minimum poly spacing	0.25
R5	Minimum gate extension of poly over active	0.22
R6	Minimum poly-active edge spacing (poly outside active area)	0.10
R7	Minimum poly-active edge spacing (poly inside active area)	0.10
R8	Minimum metal width	0.23
R9	Minimum metal spacing	0.23
R10	Poly contact size	0.22
R11	Minimum poly contact spacing	0.25
R12	Minimum poly contact to poly edge spacing	0.10
R13	Minimum poly contact to metal edge spacing	0.06
R14	Minimum poly contact to active edge spacing	0.10
R15	Active contact size	0.22
R16	Minimum active contact spacing (on the same active region)	0.22
R17	Minimum active contact to active edge spacing	0.10
R18	Minimum active contact to metal edge spacing	0.06
R19	Minimum active contact to poly edge spacing	0.16



CMOS逻辑电路

基本逻辑电路

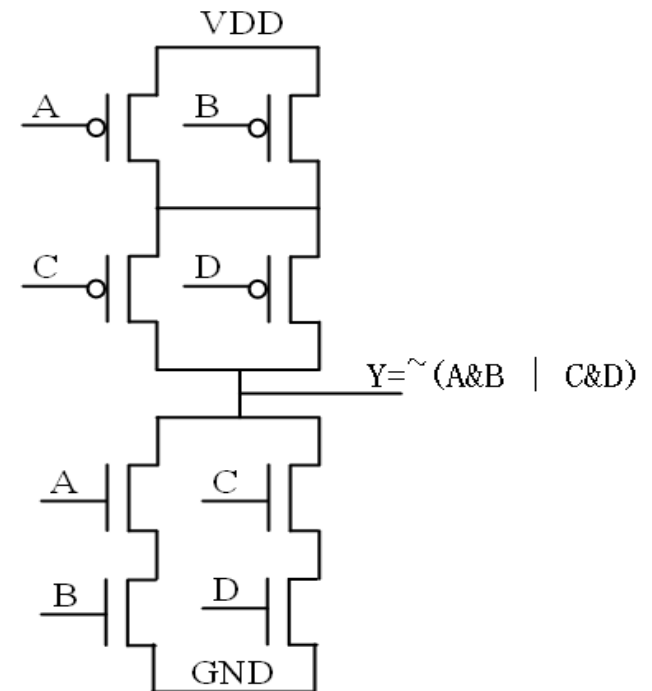
- 组合逻辑电路
 - 电路中没有存储单元，逻辑电路的输出完全由当前的输入决定。
- 时序逻辑电路
 - 电路中有存储单元，逻辑电路的输出由原来状态和当前的输入决定。

逻辑表达式

- Boolean代数的基本操作：与($*$)、或($+$)、非(\wedge)
- 一些基本定律
 - $A+0=A, A*1=A, A+1=1, A*0=0$
 - $A+A=A, A*A=A, A+\wedge A=1, A*\wedge A=0$
 - $A+B=B+A, A*B=B*A$
 - $A+(B+C)=(A+B)+C, A*(B*C)=(A*B)*C$
 - $A*(B+C)=A*B+A*C, A+(B*C)=(A+B)(A+C)$
 - $\wedge(A*B)=\wedge A+\wedge B, \wedge(A+B)=\wedge A*\wedge B$
 - $A+A*B=A, A+\wedge A*B=A+B$
- CMOS中，最基本的门是非、与非、或非门，因此逻辑表达式最好写成上述操作的组合
 - 如， $\wedge(\wedge(A*B)*\wedge(C*D))=A*B+C*D$

CMOS组合电路的组成

- 用NMOS组成正逻辑
 - 串联表示与
 - 并联表示并
- 用PMOS组成反逻辑
 - $\neg(A * B) = \neg A + \neg B$
 - $\neg(A + B) = \neg A * \neg B$
- 正反逻辑串联
 - 由于N网络导通时接地，因此输出是反向的
 - 例： $\neg(A \& B \mid C \& D)$



真值表

- 对于每一种可能的输入组合，给出输出值。对于一个 n 输入的电路，有 2^n 项。
 - 与项之或
- 以一位全加器为例
 - 输入：A, B, Cin; 输出：S, Cout
 - $S = \bar{A} * \bar{B} * C_{in} + \bar{A} * B * \bar{C}_{in} + A * \bar{B} * \bar{C}_{in} + A * B * C_{in}$
 - $C_{out} = \bar{A} * B * C_{in} + A * \bar{B} * C_{in} + A * B * \bar{C}_{in} + A * B * C_{in}$

$$= A * B + A * C_{in} + B * C_{in}$$

A	B	C in	S	C out
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

卡诺图

- 便于逻辑表达式优化的输入输出关系表示。
- 在图中相临的“1”或“X”合并成一项。
- 以一位全加器为例

		$\wedge A$		A	
		00	01	11	10
$\wedge C$	0	0	1	0	1
C	1	1	0	1	0
		$\wedge B$	B	$\wedge B$	

$$S = \wedge A * \wedge B * C_{in} + \wedge A * B * \wedge C_{in} + A * \wedge B * \wedge C_{in} + A * B * C_{in}$$

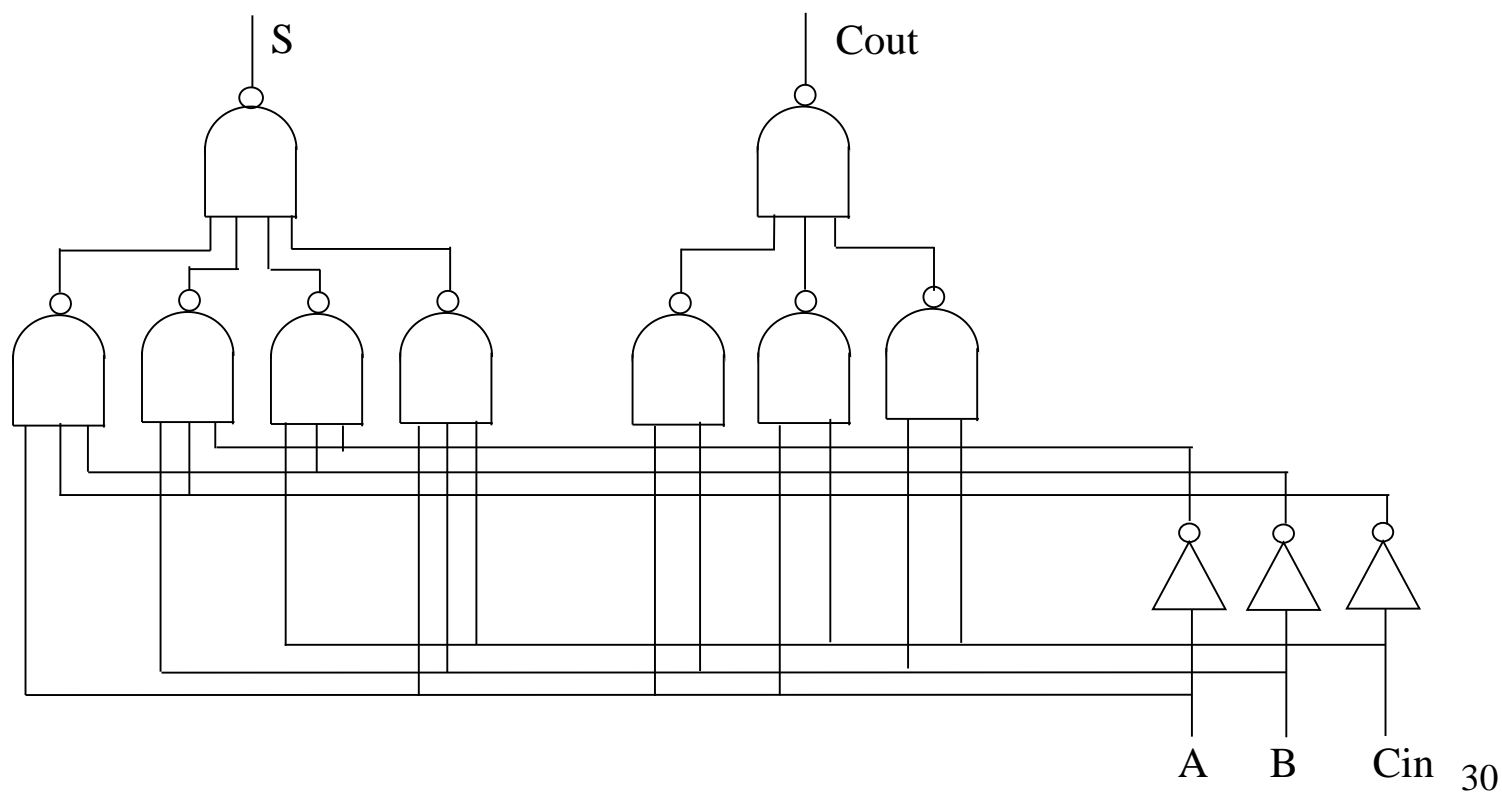
		$\wedge A$		A	
		00	01	11	10
$\wedge C$	0	0	0	1	0
C	1	0	1	1	1
		$\wedge B$	B	$\wedge B$	

$$C_{out} = A * B + A * C_{in} + B * C_{in}$$

A	B	C _{in}	S	C _{out}
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

逻辑图

- 直接用门及互连表示输入输出的逻辑关系
- 以一位全加器为例



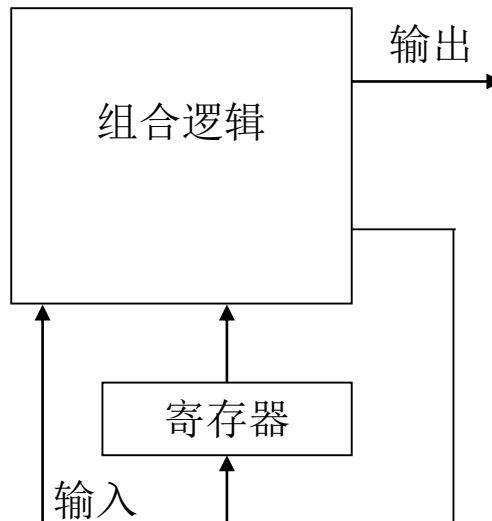
常见组合逻辑：译码器和选择器

输入			输出							
C	B	A	Y0	Y1	Y2	Y3	Y4	Y5	Y6	Y7
0	0	0	1	0	0	0	0	0	0	0
0	0	1	0	1	0	0	0	0	0	0
0	1	0	0	0	1	0	0	0	0	0
0	1	1	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0
1	0	1	0	0	0	0	0	1	0	0
1	1	0	0	0	0	0	0	0	1	0
1	1	1	0	0	0	0	0	0	0	1

输入			输出
C	B	A	Y
0	0	0	D0
0	0	1	D1
0	1	0	D2
0	1	1	D3
1	0	0	D4
1	0	1	D5
1	1	0	D6
1	1	1	D7

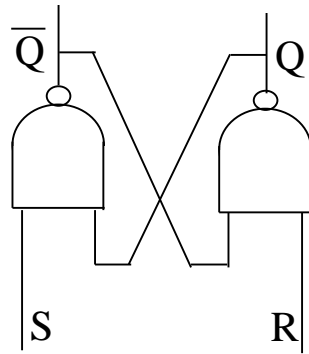
时序逻辑电路

- 时序逻辑电路内部有存储单元，其行为由输入和内部单元的值共同决定
 - 可分为同步时序逻辑电路和异步时序逻辑电路，计算机中主要用同步电路
 - 在同步时序电路中，所有存储单元的变化由时钟统一触发

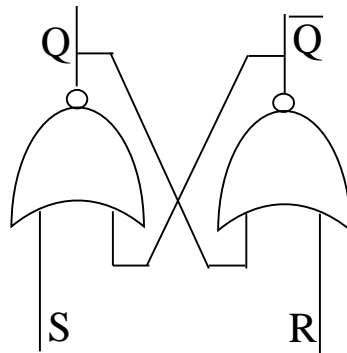


RS触发器

- 其它寄存器电路的基础



R	S	Q	$\sim Q$
0	0	nc	nc
0	1	1	0
1	0	0	1
1	1	Q	$\sim Q$

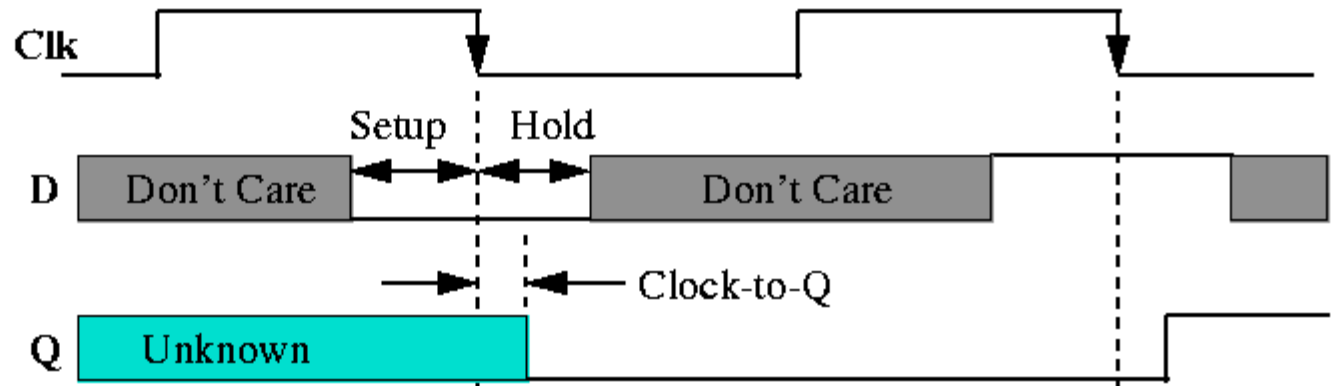
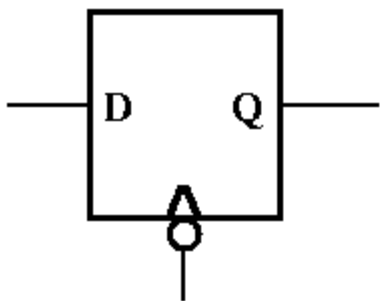
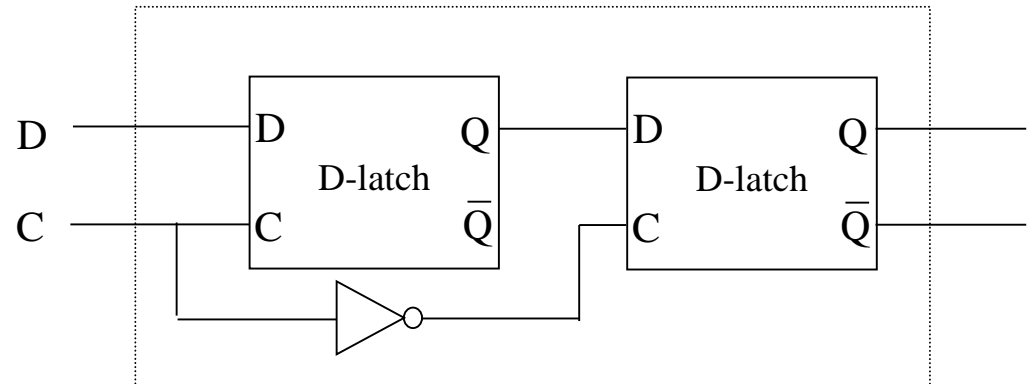
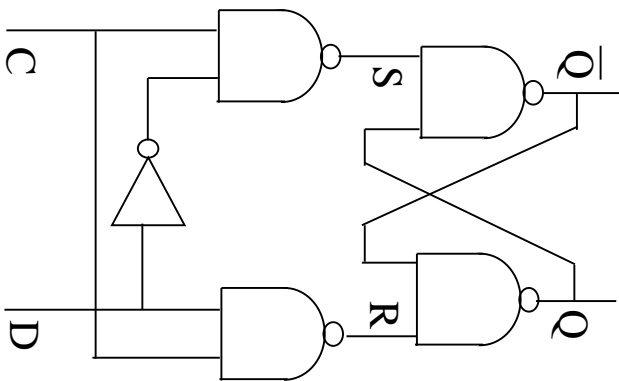


R	S	Q	$\sim Q$
0	0	Q	$\sim Q$
0	1	1	0
1	0	0	1
1	1	nc	nc

D门闩和D触发器

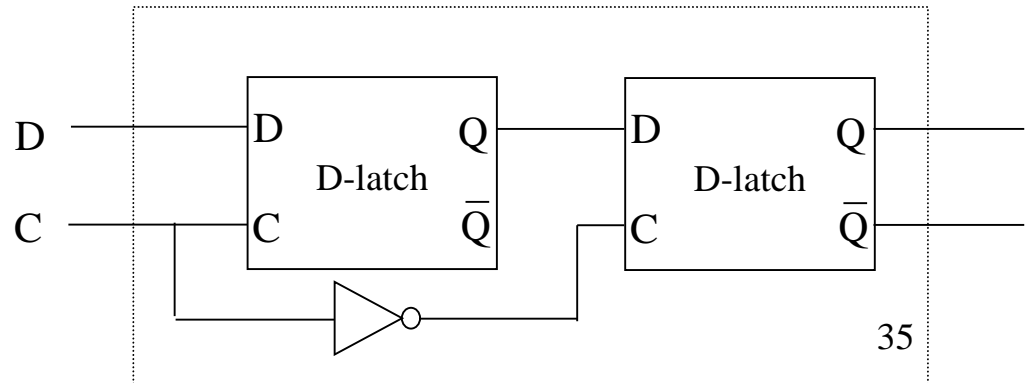
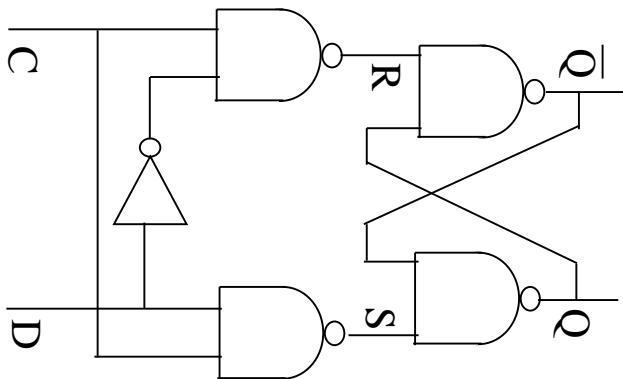
- **D门闩和D触发器**

- 三个重要时间参数: **Setup**、**Hold**、**CLK-to-Q**
- 触发器的延迟=**Setup+CLK-to-Q**



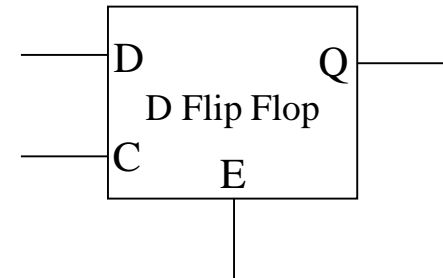
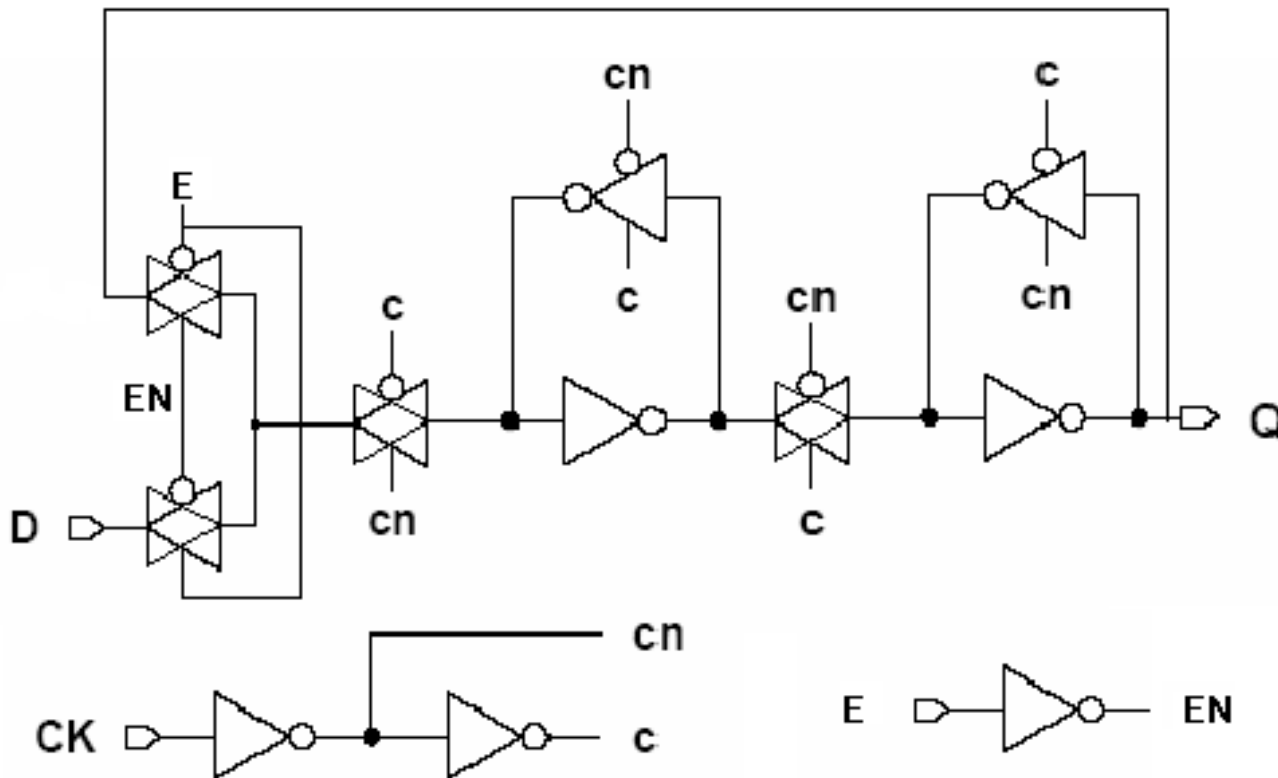
D触发器时间参数计算

- **Hold:** 时钟下降后数据变化不影响第一级D-Latch
 - 在本电路中为0
- **Setup:** 时钟沿下降之前保证数据已采到第一级D-latch
 - 1级反相器 + 3级与非门 (Q端从1变0)
- **Clock-to-Q:** 时钟沿下降后引起Q的变化
 - 1级反相器 (C反相器与D方向器并行) + 3级与非门
- 上述只是初略计算, 具体需要仿真
 - Setup和CLK-to-Q的“划分” (可能D不用到达一级门闩的Q时钟沿就变化也可以, 但总延迟不变); 还有连线延迟影响
- 如果C先通过两级反相器驱动再用, 对上述三个时间有何影响?



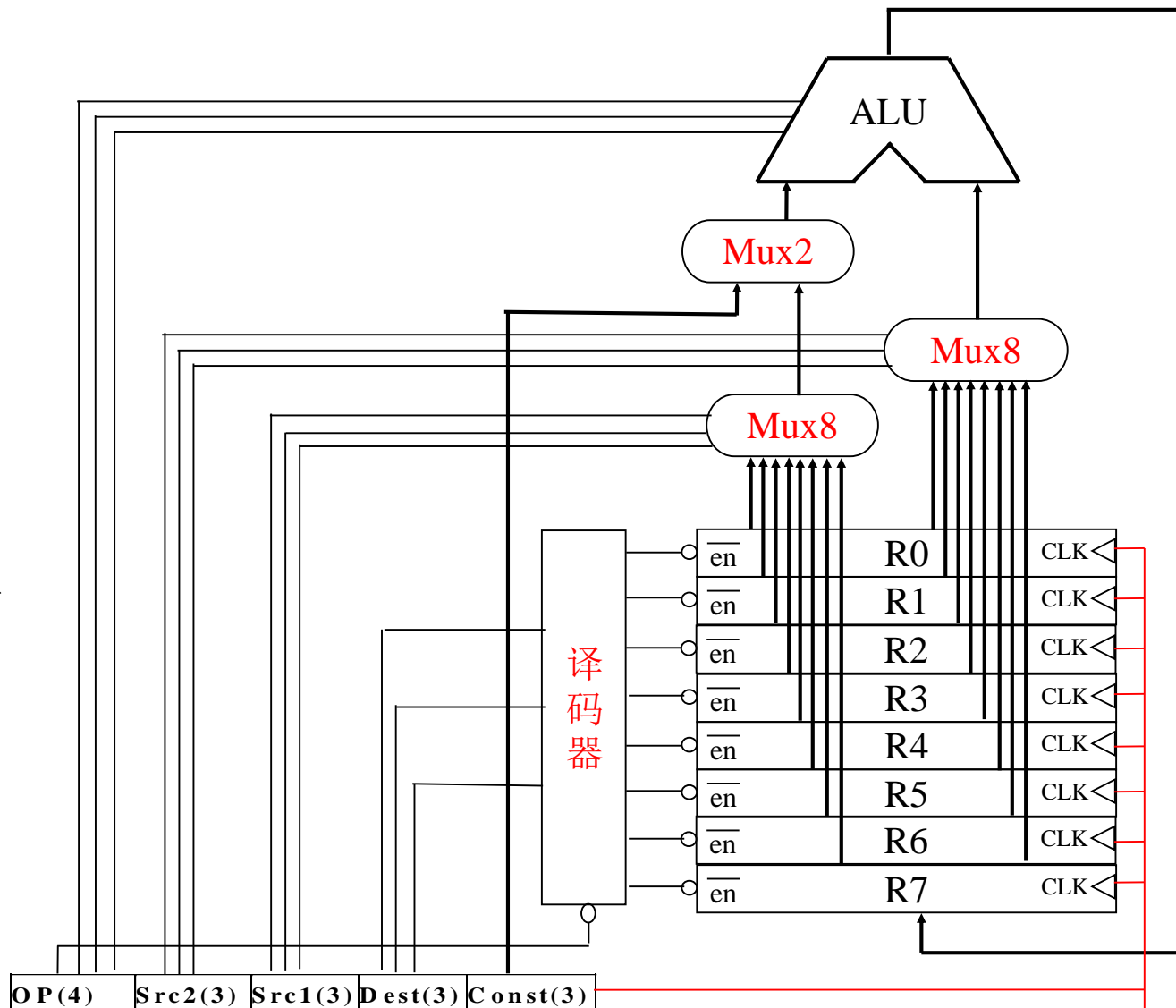
EDFF电路

- 帶輸入使能的D触发器
 - 只有E有效，才能输入



例：简单CPU的运算电路

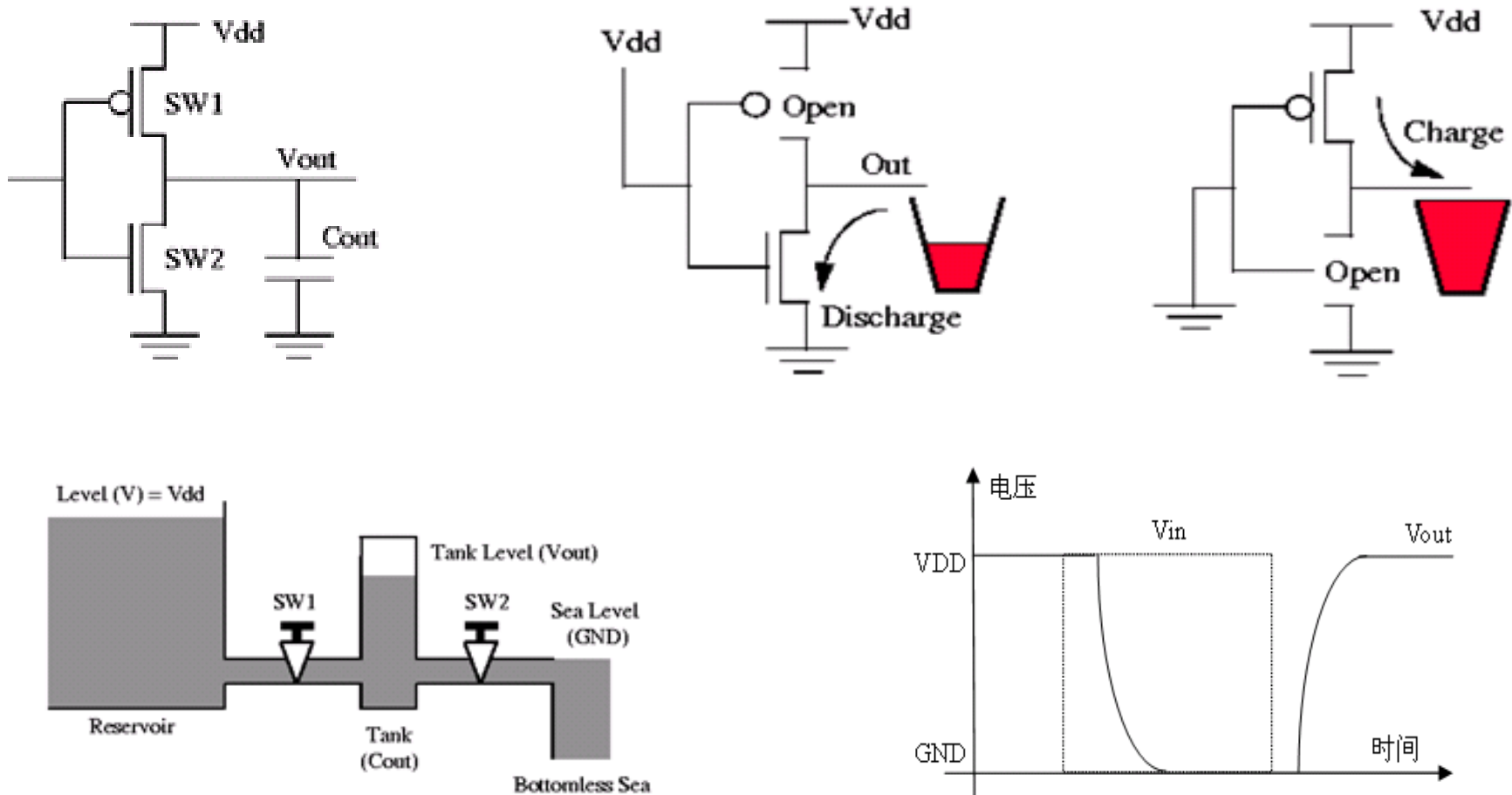
- **16位IR**
- **8个16位GPR**
- **4位OP**
- **三地址**
- **Load-store结构**



CMOS电路延迟

CMOS电路延迟原理

- 输入电压变化时，输出电压的变化不是瞬时完成的



CMOS电路延迟模型

- 与延迟相关的因素
 - 输入transition、内部延迟、输出负载有关
 - 负载包括下一级负载、连线延迟等
 - 内部延迟(ns)+与负载（包括负载）有关的延迟（ns/fF）
- 二输入与非门为例

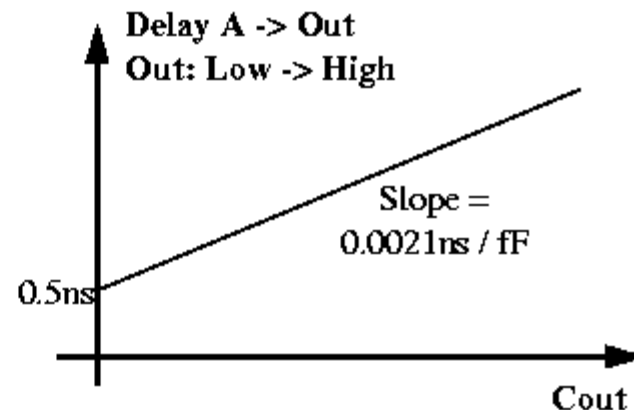


For A and B: Input Load = 61 fF

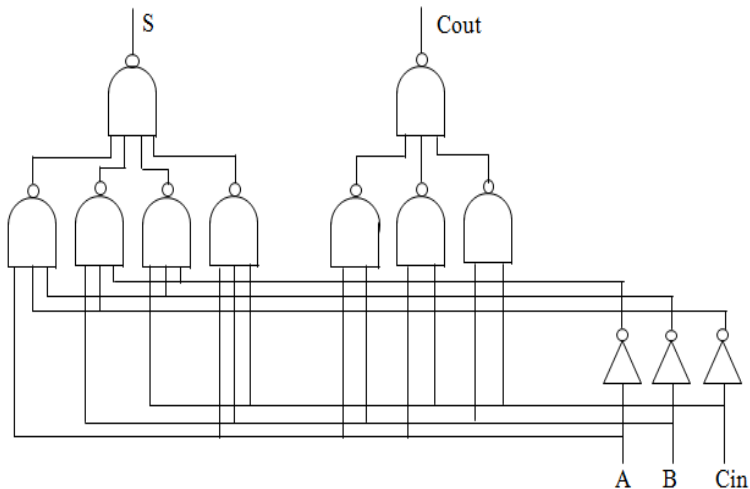
For either A -> Out or B -> Out:

$T_{Plh} = 0.5\text{ns}$ $T_{plhf} = 0.0021\text{ns} / \text{fF}$

$T_{Phl} = 0.1\text{ns}$ $T_{Phlf} = 0.0020\text{ns} / \text{fF}$



CMOS电路延迟例子



$$T_{Ph1} (INV) = 0.10ns$$

$$TP1h(INV) = 0.10ns$$

$$T_{Ph1}(NAND2) = 0.10ns$$

$$TP1h(NAND2) = 0.20ns$$

$$T_{Ph1}(NAND3) = 0.15ns$$

$$TP1h(NAND3) = 0.20ns$$

$$T_{lhf}(\text{INV}) = 0.010\text{ns/fF}$$

$$\text{Th1f (INV)} = 0.010\text{ns/fF}$$

$$T_{1hf}(\text{NAND2}) = 0.010\text{ns/fF}$$

$$\text{Th1f (NAND2)} = 0.015\text{ns/fF}$$

$$T_{1hf}(\text{NAND3}) = 0.010\text{ns/fF}$$

$$\text{Th1f (NAND3)} = 0.020\text{ns/fF}$$

$$C_{il}(INV) = 20\text{fF}$$

$$C_{il}(NAND2/3) = 30fF$$

$$C_{\text{wire}} = 50\text{fF}$$

- 假设Cout负载为60fF.

$$\begin{aligned} T_{1h}(C_{in} \rightarrow C_{out}) &= T_{Ph1}(NAND2) + T_{1hf}(NAND2) * (C_{i1}(NAND3) + C_{wire}) + \\ &\quad T_{P1h}(NAND3) + T_{1hf}(NAND3) * C_{o1} \\ &= 0.10 + 0.015 * (30 + 50) + 0.10 + 0.010 * 60 \\ &= 2.0ns \end{aligned}$$

标准单元延迟的计算

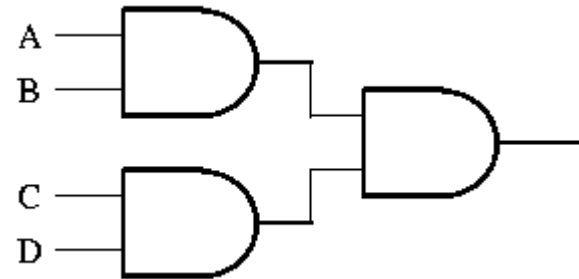
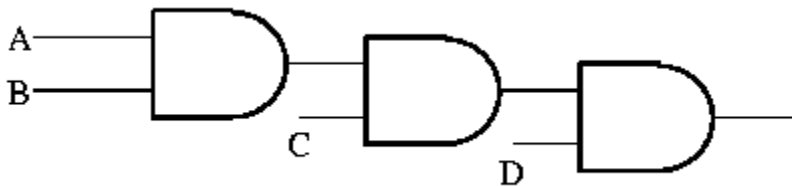
- 深亚微米电路中连线延迟占总延迟的大部分
 - 以TSMC18的NAND2X1为例：内部延迟0.034ns，每pf延迟为4.5ns左右，输入电容0.004pf
- 连线延迟的计算是一个不断迭代求精的过程，
 - 综合时根据负载个数估计线长，非常初略
 - 布局后根据距离估计线长，比较准确了
 - 布线后进一步考虑具体连线、互相干扰、过孔等精确连线信息
- TSMC 0.18微米二输入与门的特征
 - 面积、输入电容、最大负载、功耗
 - 延迟（非线性）：根据输入transition和输出负载查表
 - [库单元的定义](#)

降低延迟的方法

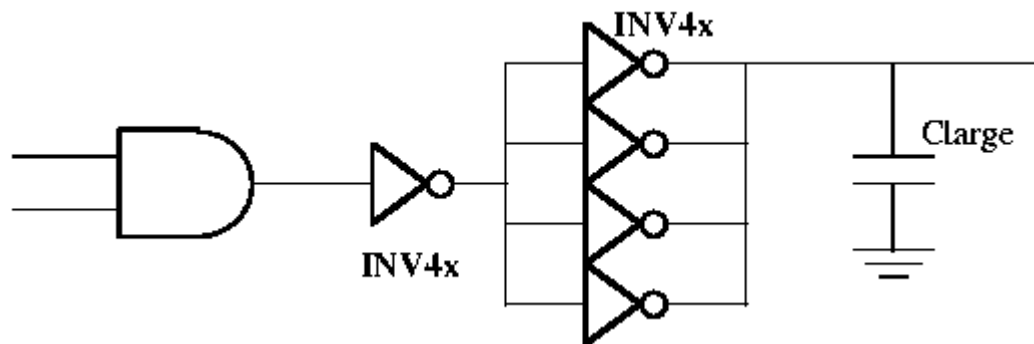
- 通过结构设计降低延迟
 - 如流水级的划分、**CACHE**读出和命中比较是否在同一级、浮点流水级设计、**Pipeline stalling signal** 的设置
- 逻辑设计
 - 如多选一通路的选择次序、加法器算法、流水级间的局部调整等
- 物理设计
 - 动态电路、**clock skew**、**clock jitter**、**transition time**、布局布线

通过逻辑设计降低延迟

- 减少门的级数



- 负载平衡: CMOS的负载能力较低, 分推



从Verilog到版图

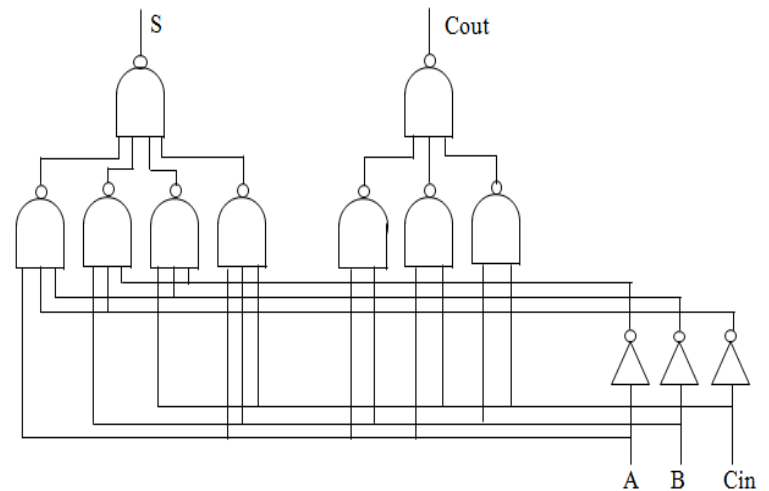
Verilog语言

- 做结构设计时要明白逻辑上是什么样子
- 做逻辑设计时要明白物理上是什么样子
- 硬件设计人员和软件设计人员对Verilog有不同认识
- 我只用Verilog中三种语句
 - 组合逻辑: `assign`
 - 时序逻辑: `always (@posedge clock)`
 - 模块调用

Verilog 语言---组合逻辑

- 组合逻辑---assign语句
 - 以一位全加器为例

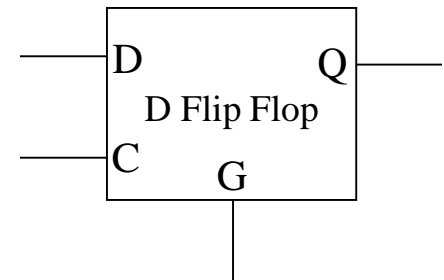
```
module full_adder(a,b,c,s,cout);  
    input a,b,c;  
    output s,out;  
  
    assign s=a^b^c;  
    assign cout=a&b | b&c | c&a  
endmodule
```



Verilog语言----时序逻辑

- 时序逻辑---触发器
 - 以一位触发器为例

```
module dflipflop(D,C,G,Q);  
input D,C,G;  
output Q;  
reg Q;  
always (@posedge C)begin  
    if (G) Q<=D;  
end  
endmodule
```



Verilog语言----模块调用

- 以四位串行加法器为例

```
module adder4(a,b,cin,s,cout);  
  input [3:0] a;  
  input [3:0] b;  
  input cin;  
  output [3:0] s;  
  output cout;  
  wire carry[2:0]  
  full_adder bit0(.a(a[0]),.b(b[0]),.c(cin[0]),.s(s[0]),.cout(carry[0]));  
  full_adder bit1(.a(a[1]),.b(b[1]),.c(carry[0]),.s(s[1]),.cout(carry[1]));  
  full_adder bit2(.a(a[2]),.b(b[2]),.c(carry[1]),.s(s[2]),.cout(carry[2]));  
  full_adder bit3(.a(a[3]),.b(b[3]),.c(carry[2]),.s(s[3]),.cout(cout));  
endmodule
```

Verilog语言

- 四位计数器

```
module counter4(reset,en,clock,out);  
input reset,en,clock;  
output [3:0] out;  
reg [3:0] counter;  
wire [3:0] c_in;  
assign out = counter;  
adder4 counter(.a(counter),.b(4'b0),.cin(1'b1),.s(c_in),.cout());  
always (@posedge clock) begin  
counter <= reset ? 0 : en ? c_in : counter;  
//counter<=reset ? 0 : en ? counter +1 : counter;  
end  
endmodule
```

从Verilog 到GDSII

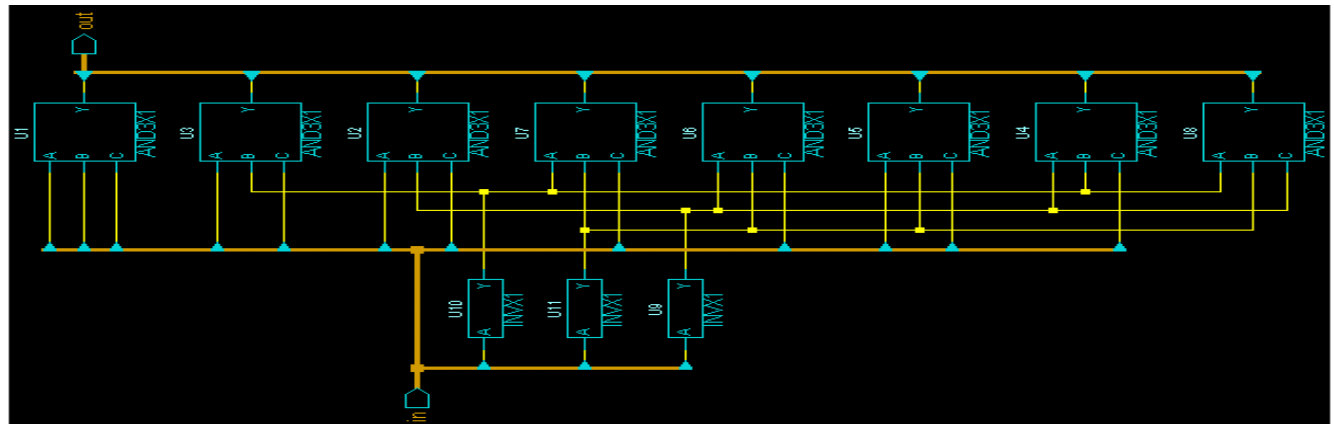
- 以3-8译码器为例
 - Verilog
 - 网表
 - 布线
 - GDSII

Verilog

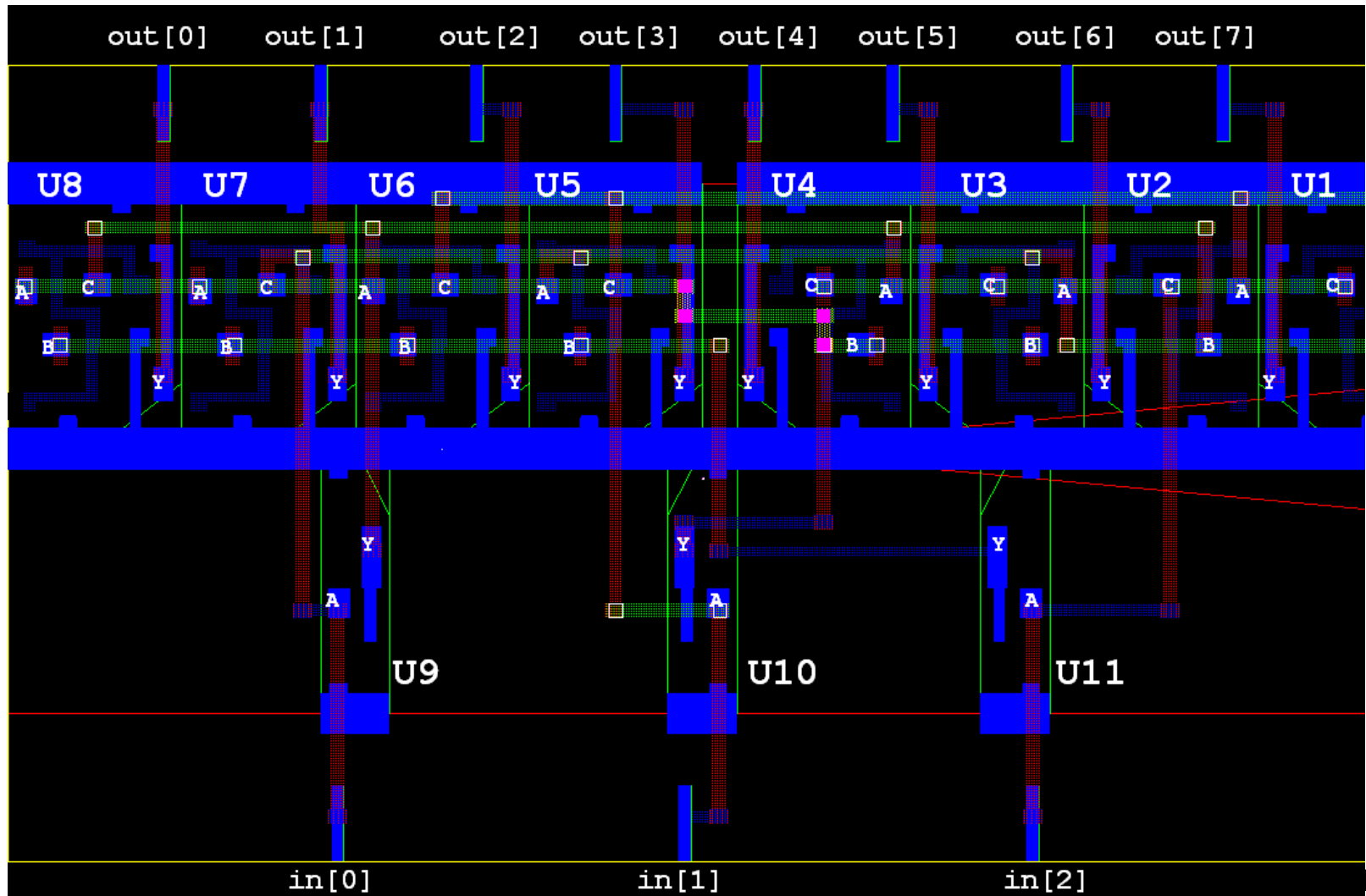
```
module decode(in, out);  
    input [2:0] in;  
    output [7:0] out;  
  
    assign out[0] = (!in[2]) && (!in [1]) && (!in[0]); // (in==3'b000)?1:0;  
    assign out[1] = (!in[2]) && (!in [1]) && ( in[0]); // (in==3'b001)?1:0;  
    assign out[2] = (!in[2]) && ( in [1]) && (!in[0]); // (in==3'b010)?1:0;  
    assign out[3] = (!in[2]) && ( in [1]) && ( in[0]); // (in==3'b011)?1:0;  
    assign out[4] = ( in[2]) && (!in [1]) && (!in[0]); // (in==3'b100)?1:0;  
    assign out[5] = ( in[2]) && (!in [1]) && ( in[0]); // (in==3'b101)?1:0;  
    assign out[6] = ( in[2]) && ( in [1]) && (!in[0]); // (in==3'b110)?1:0;  
    assign out[7] = ( in[2]) && ( in [1]) && ( in[0]); // (in==3'b111)?1:0;  
  
endmodule
```

网表

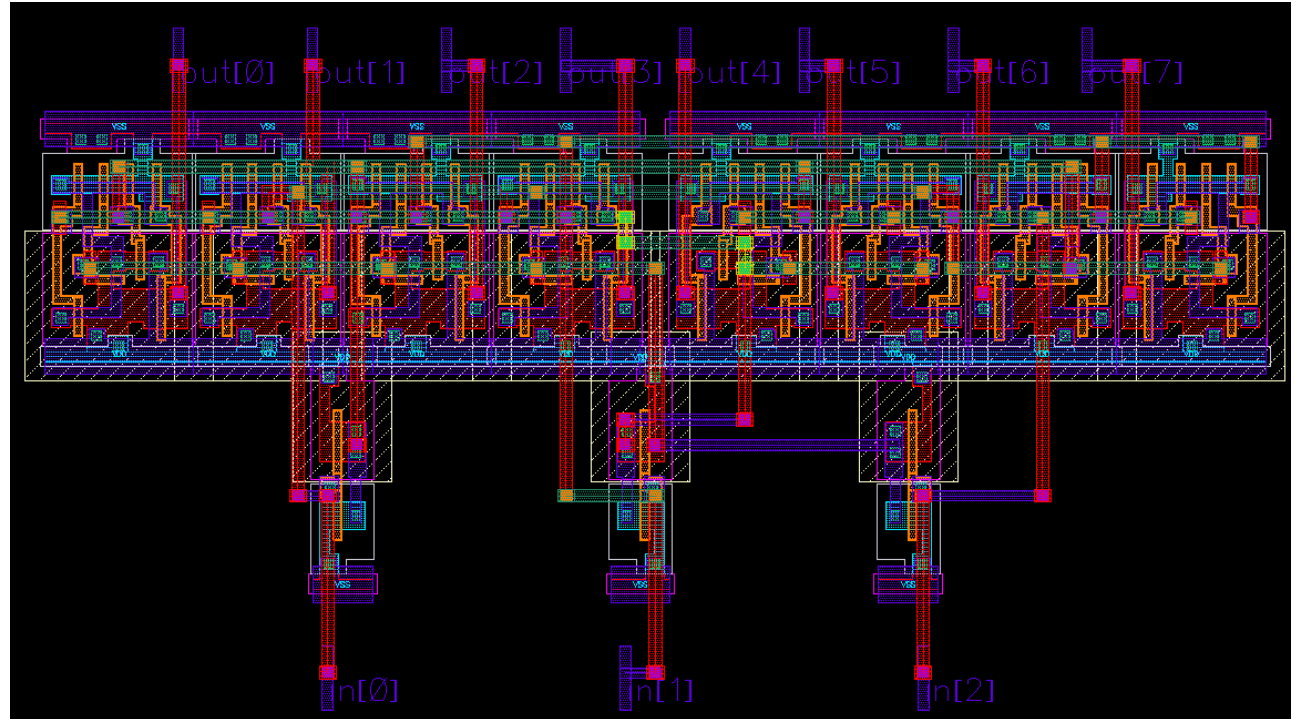
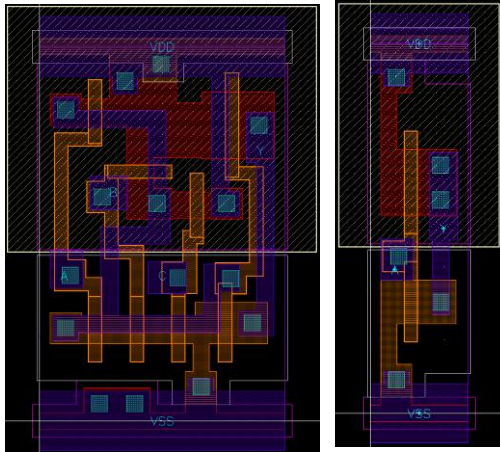
```
module decode ( in, out );
input  [2:0] in;
output [7:0] out;
    wire n1, n2, n3;
    AND3X1 U1 ( .A(in[1]), .B(in[0]), .C(in[2]), .Y(out[7]) );
    AND3X1 U2 ( .A(in[1]), .B(n1), .C(in[2]), .Y(out[6]) );
    AND3X1 U3 ( .A(in[0]), .B(n2), .C(in[2]), .Y(out[5]) );
    AND3X1 U4 ( .A(n1), .B(n2), .C(in[2]), .Y(out[4]) );
    AND3X1 U5 ( .A(in[0]), .B(n3), .C(in[1]), .Y(out[3]) );
    AND3X1 U6 ( .A(n1), .B(n3), .C(in[1]), .Y(out[2]) );
    AND3X1 U7 ( .A(n2), .B(n3), .C(in[0]), .Y(out[1]) );
    AND3X1 U8 ( .A(n2), .B(n3), .C(n1), .Y(out[0]) );
    INVX1 U9 ( .A(in[0]), .Y(n1) );
    INVX1 U10 ( .A(in[1]), .Y(n2) );
    INVX1 U11 ( .A(in[2]), .Y(n3) );
endmodule
```



布线后



版图



其它 “0”和 “1”的表示方法

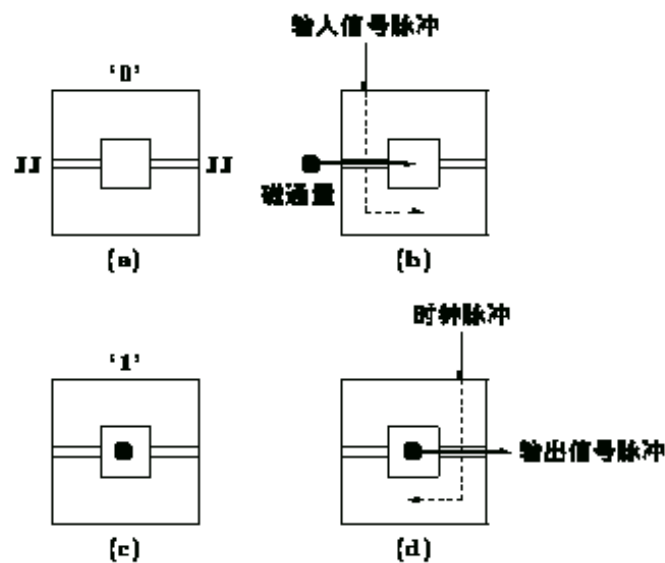
其它 “0”和 “1”表示方法

- 用磁通量的有无表示，超导体工艺
- 用能级的高低表示，量子计算机
- 用基因序列表示，A, G, C, T, DNA计算机，非二进制？

超导计算---RSFQ技术

(Rapid Single Flux Quantum)

- 基本原理
 - 超导 (4-5K) 环中的磁通量具有量子化特性，设计电路使超导环中的磁通量只变化一个磁通量，用磁通量的有无来表示二进制数位的“1”和“0”。磁通量的变化由外加电流控制。
 - SQUID (Superconducting Quantum Interface Device)
- 电路特点
 - 工作频率高——100 GHz (实验室已达370 GHz, 1.5 μ 工艺)。
 - 每个门的功率——0.1 μ W。
 - 工艺比较简单。



量子计算（1）

- 量子力学特性

- 量子叠加态(superposed state): 量子器件的信息位称为量子位(qubit), 它可处于叠加态。叠加态可以是 “0” 也可以是 “1”。通过测量或与其他物体发生相互作用可呈现出 “0” 态或 “1” 态。由于每个量子位都可以是 “0” 或 “1”, n 个量子位就可以表示 2^n 个 n 位数。常规计算机的一个 n 位存储单元只能存放一个 n 位数, 而 n 个量子位可以存放 2^n 个 n 位数, 可以实现超大容量的存储器。
- 量子纠缠态(entangled state)——除了叠加态以外, 用作运算的多个量子位还应处于纠缠态, 即所有量子位的状态紧密相关。当测量某个量子位时, 会影响其他量子位的测量结果。
- 量子并行 (quantum parallelism)——计算 $f(x)$ 时, 可同时计算出 x 的所有值的 $f(x)$ 。所以不需要多次循环, 也不需要多个处理机并行计算。

量子计算（2）

- 应用领域

- 大数 N 因子分解——令 $n = \log_2 N$ ，经典算法所需步骤为 $2^{n/2}$ ，Shor 量子并行算法所需步骤为 $\text{Poly}(n)$ ， $\text{Poly}(n)$ 为 n 的多项式。该算法将 NP 问题转换为 P 问题。
- 搜寻算法——在 N 个元素的集合中搜寻某个元素，经典算法搜寻 $N/2$ 次后，找到的概率为 $1/2$ ，Grover 量子搜寻算法则只需 $N^{1/2}$ 次，即可达到同样概率。
- 量子系统模拟——常规计算机不可能有效地模拟量子系统，因为它们的物理机制不同。用常规计算机模拟量子系统，所需的信息量和时间都远大于模拟经典系统。量子计算可用于研究高温高密度等离子体、量子色动力学、晶体固态模型、分子行为的量子模型等。

量子计算（3）

- 物理实现

- 量子位的实现——任何两态的量子系统都可作为量子位，如原子的能级、电子或原子核的自旋、光子的正交偏振态等。
- 量子计算机的实现——有多种可能，包括：核磁共振（用磁场中的原子核自旋作为量子位）、离子阱（用被俘获在线性量子阱中的离子作为量子位）、硅基半导体量子器件（杂质核自旋与电子自旋相互作用）等。

- 存在问题

- 量子位的缠结态容易崩溃，位数越多，越难实现。
- 量子器件之间的连接。
- 为了维持量子逻辑的一致性，量子系统和环境的隔离。
- 设备缺陷所引起的逻辑错误。

分子计算机

- 原理

- DNA 计算机利用 DNA 分子保存信息。DNA 分子是由 A, G, C, T 四种核苷酸（碱基）组成的序列。不同的序列可用来表示不同的信息。通过 DNA 分子之间的一系列生化反应来进行运算，可产生表示结果的 DNA 分子。已解决了 7 个城市的旅行售货员等问题。

- 优点

- 高度并行——所有 DNA 分子同时运算。
- 能耗低——半导体计算机的 10^{10} 之一。
- 存储密度大——磁存储器的 10^{12} 倍。

- 缺点

- 生化反应慢、操作有随机性、DNA 分子容易水解、DNA 分子之间难以通信。

本章小结

- 二进制与逻辑电路是本科已经学习的内容，硕士的要求比本科深入一些
 - 浮点格式与数值转换
 - 简单的CMOS晶体管级数字电路设计
 - CMOS电路的延迟计算
 - 触发器原理和 Setup、Hold、CLK-to-Q时间计算
 - Verilog语言

作业

- 书后不带“*”的习题
- 第一次作业补充两道题，可以随第二次作业交
 - 在个人电脑上运行 SPEC CPU2000程序，分别给出无编译选项优化和-O2优化的运行结果
 - 分别在苹果手机、华为手机及 Windows机器上测试Octane分值，并简单评述（<http://chromium.github.io/octane>）