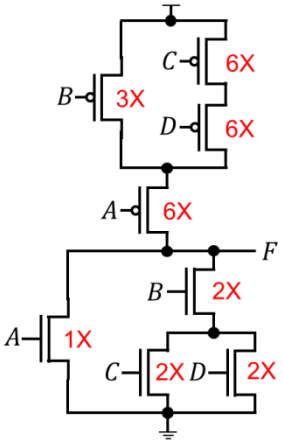
Homework Assignment 4
(Due Feb. 17th at the beginning of the class)

- R_n is the resistance of a 1X nFET.
- $\mu_n = 2\mu_p$.
- C_L is the load capacitance connected to the output.
- Target time constant: $R_n C_L$

(1) **[FET Sizing, 10 points]** Size the transistors in the following gate to satisfy the time constant. **Try to minimize** the total width of the transistors.



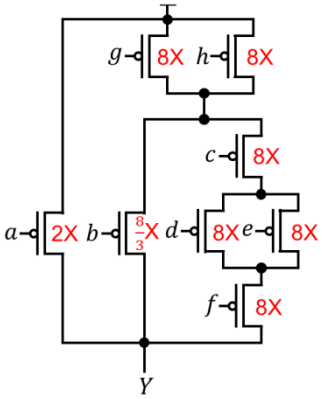
Total width = 7 + 21 = 28X

The best solution is as follows:

- For the nFET network, B=bX, C=D=cX. Then, $\frac{R_n}{b} + \frac{R_n}{c} = R_n$, which is $\frac{1}{b} + \frac{1}{c} = 1$. Minimize $b + 2c$. $c = \frac{b}{b-1}$. $f(b,c) = b + 2c = f(b) = b + \frac{2b}{b-1}$. $f'(b) = 1 + \frac{2(b-1)-2b}{(b-1)^2} = 1 - \frac{2}{(b-1)^2} = 0 => b = 1 + \sqrt{2}, c = 1 + \frac{1}{\sqrt{2}}$
- For the pFET network, B=bX, C=D=2bX, A=cX. Then, $\frac{2R_n}{b} + \frac{2R_n}{c} = R_n$, which is $\frac{1}{b} + \frac{1}{c} = \frac{1}{2}$. Minimize $5b + c$. $c = \frac{2b}{b-2}$. $f(b,c) = 5b + c = f(b) = 5b + \frac{2b}{b-2}$. $f'(b) = 5 + \frac{2(b-2)-2b}{(b-2)^2} = 5 - \frac{4}{(b-2)^2} = 0 => b = 2 + \frac{2}{\sqrt{5}}, c = 2 + 2\sqrt{5}$

Total width = $(1+\sqrt{2})+2\left(1+\frac{1}{\sqrt{2}}\right)+5\left(2+\frac{2}{\sqrt{5}}\right)+(2+2\sqrt{5})=15+2\sqrt{2}+4\sqrt{5}\approx 26.77X$

(2) **[FET Sizing, 10 points]** Size the transistors in the following pFET network to satisfy the time constant. **<u>Try to minimize</u>** the total width of the transistors.



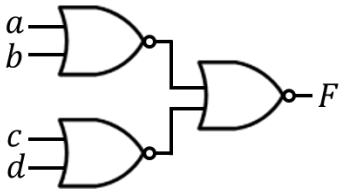
Total width = $(50+\frac{8}{3})X \approx 52.67X$.

(If you minimize the path "g-b", first, g=h=b=4X. Then, c=d=e=f=12X. In this case, the total width = 2X + 3*4X + 4*12X = 62X.)

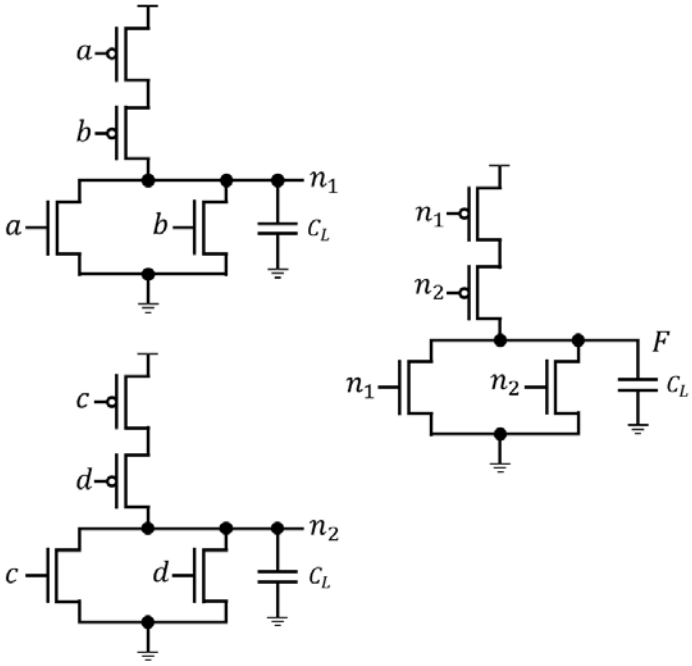
(3) **[FET Sizing, 10 points]** We are going to design the following logic:

$$\boldsymbol{F=(a+b)\cdot(c+d)}$$

We can design this logic as follows:

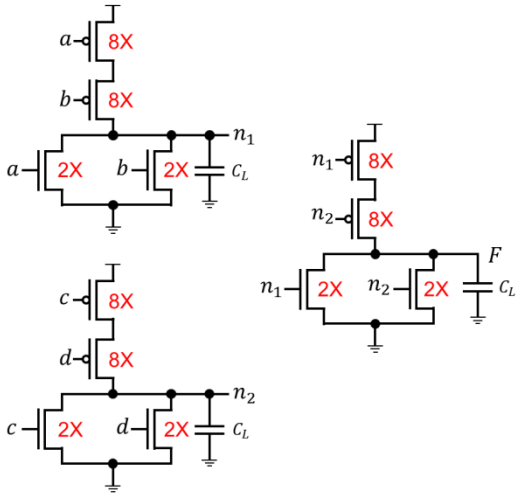


The schematic shown above can be modeled as follows to compute the delay from the inputs to the output:



To simplify the problem, we assume that the delay from an input to the output is computed by $\tau_1 + \tau_2$ where τ_1 and τ_2 are the time constants for charging or discharging the first and the second NOR gates, respectively. The target time constant is still $R_n C_L$. Size all the transistors in the three NOR gates to satisfy the target time constant. **Minimize** the total width of the transistors.

1) Assuming $\tau_1 = \tau_2 = \frac{R_n C_L}{2}$, we get the following solution:



In this case, the total width is 60X.

2) However, we intuitively know that the first-stage NOR gates are the bottleneck for the area minimization. Suppose the first-stage NOR gates are aX NOR gates and the second-stage NOR gate is a bX NOR gate. Then, the delay of the first-stage NOR gate is $\frac{R_n C_L}{a}$ and that of the second-stage NOR gate is $\frac{R_n C_L}{b}$. Satisfying the total delay leads to $\frac{R_n C_L}{a} + \frac{R_n C_L}{b} = R_n C_L$, so $\frac{1}{a} + \frac{1}{b} = 1$. The total area is $(2a + b)$*(area of a 1X NOR gate). Minimize $2a + b$ when $b = \frac{a}{a-1}$. $f(a) = 2a + \frac{a}{a-1}$. $f'(a) = 2 + \frac{(a-1)-a}{(a-1)^2} = 2 - \frac{1}{(a-1)^2} = 0 \Rightarrow a = 1 + \frac{1}{\sqrt{2}}, b = 1 + \sqrt{2}$. Thus, the first-stage NOR gates should be $(1 + \frac{1}{\sqrt{2}})$X and the second-stage NOR gate should be $(1 + \sqrt{2})$X.

In this case, the total width = $10*\{(1 + \frac{1}{\sqrt{2}}) * 2 + (1 + \sqrt{2})\} = 30 + 20\sqrt{2} \approx 58$X.

3) In addition, we can consider different charging/discharging delays. Suppose the sizes of the nFETs and pFETs in the first-stage NORs are aX and bX, respectively, and those in the second-stage NOR are cX and dX, respectively. If the outputs of the first-stage NORs are 0 (the nFETs are turned on), the output of the second-stage NOR is 1 (the pFETs are turned on). Then, $\tau_1 = \frac{R_n C_L}{a}$ and $\tau_2 = \frac{4R_n C_L}{d}$, so we obtain $\frac{1}{a} + \frac{4}{d} \leq 1$. Similarly, if the output of only one of the first-stage NORs is 1 (the pFETs are turned on), the output of the second-stage NOR is 0 (one of its nFETs is turned on). Then, $\tau_1 = \frac{4R_n C_L}{b}$ and $\tau_2 = \frac{R_n C_L}{c}$, so we obtain $\frac{4}{b} + \frac{1}{c} \leq 1$. The total width = $4a + 4b + 2c + 2d$. Since the two terms are independent of each other, we minimize $4a + 2d$ and $4b + 2c$ separately. For $4a + 2d$, we obtain $a = (1 + \sqrt{2})X$ and $d = (4 + 2\sqrt{2})X$. For $4b + 2c$, we obtain $b = (4 + \sqrt{2})X$ and $c = (1 + 2\sqrt{2})X$.

The total width = $(30 + 16\sqrt{2})$X ≈ 52.63X.

4) Furthermore, we might not need to size the two first-stage NOR gates equally. If you derive timing constraint inequalities for that, however, equal sizing of them gives us the best solution.