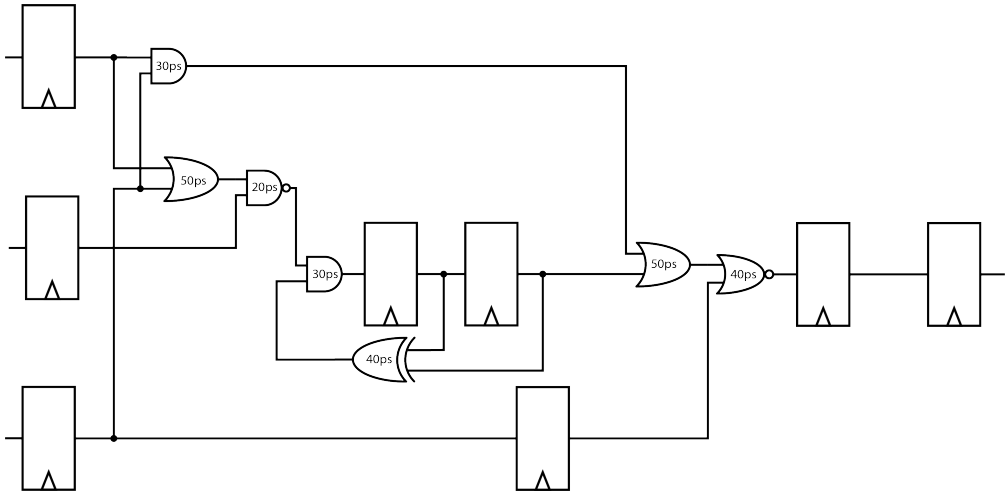
# EECS 151/251A Homework 6

Due Monday, Mar 8th, 2021

**Remember, please include a short (1-2 sentence) explanation with each answer unless otherwise directed in the problem.**

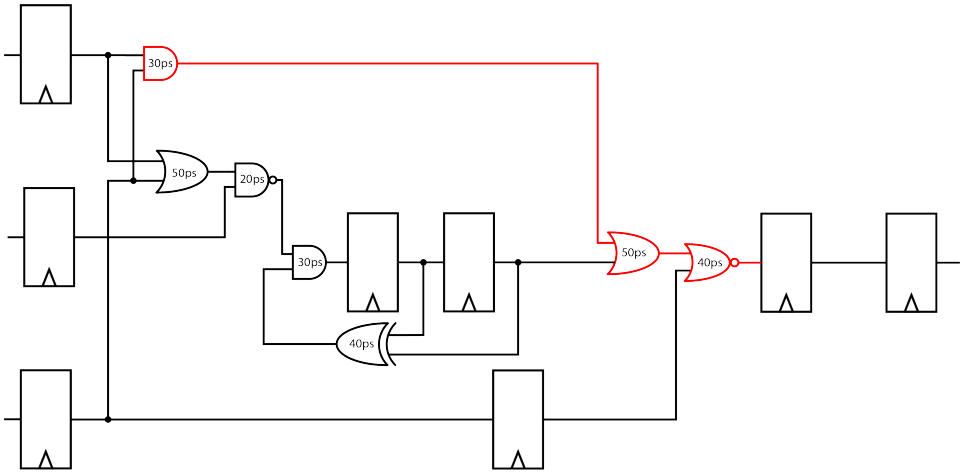
## Problem 1: Retiming

Consider the circuit below. What is the critical path delay? Retime the circuit to minimize the critical path delay. What is the maximum clock frequency for the retimed circuit? The flip-flops have $t_{\text{setup}} = 20\,\text{ps}$, $t_{\text{clk-q}} = 10\,\text{ps}$, and $t_{\text{hold}} = 5\,\text{ps}$.

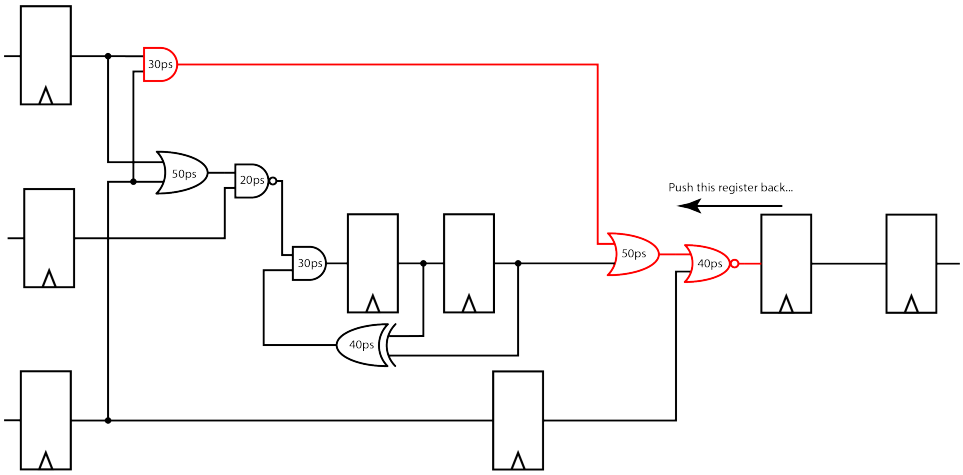


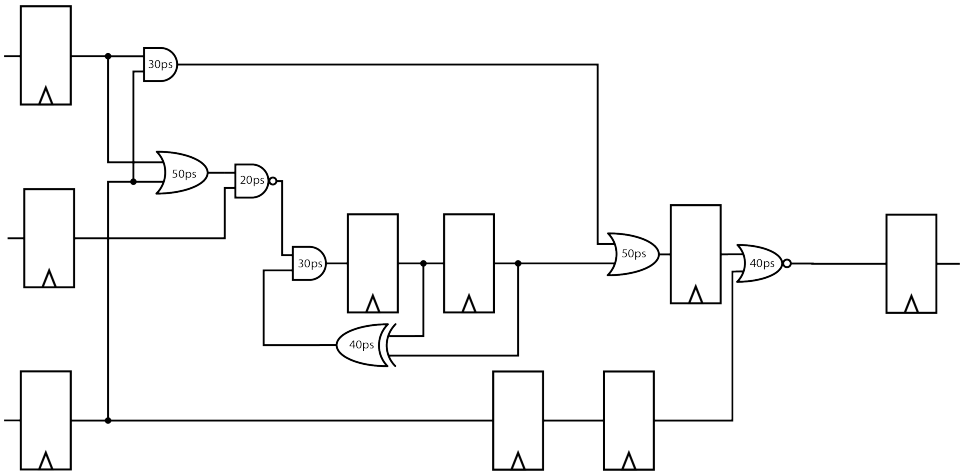

**Solution:**

The critical path of this circuit is 120 ps long, as marked in red.



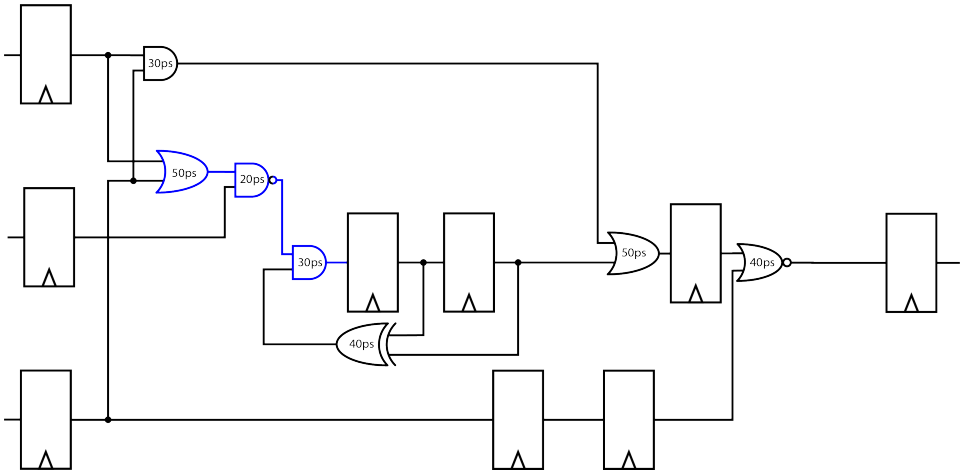

To break this path, we will need to push the right register back through the NOR.



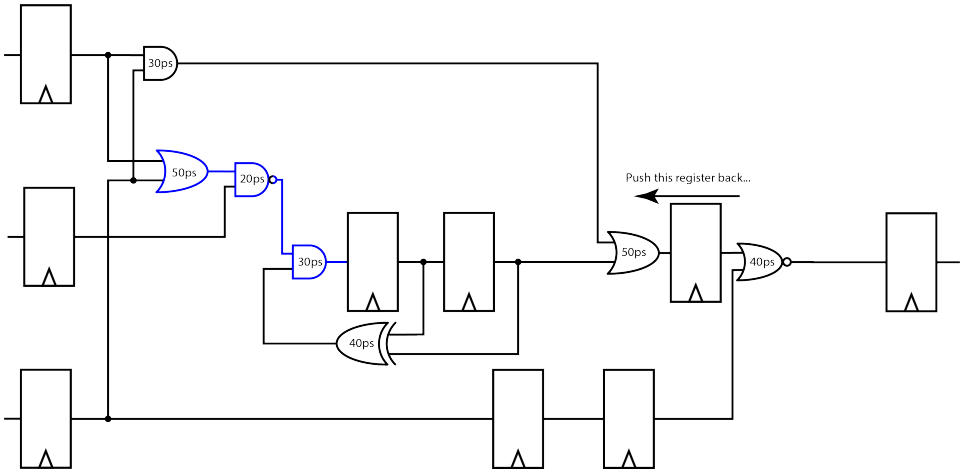



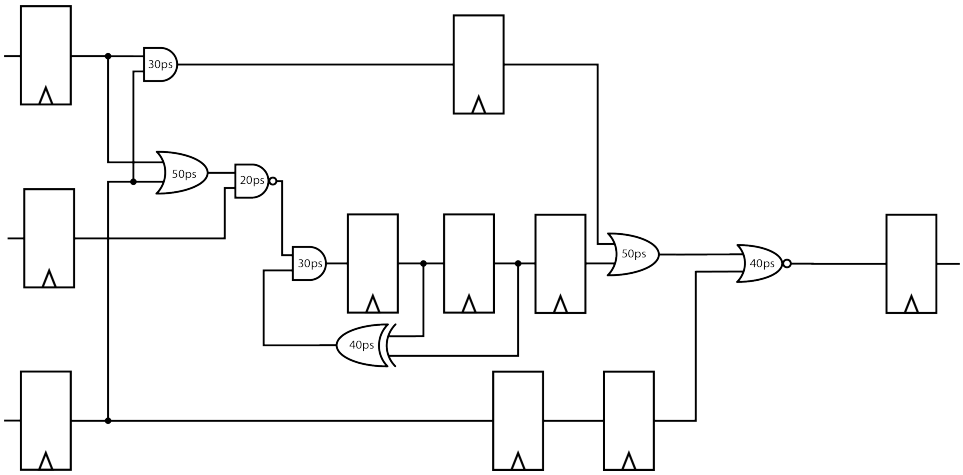

However, now that this path is no longer critical, the new critical path is the input to the loop,

marked here in blue

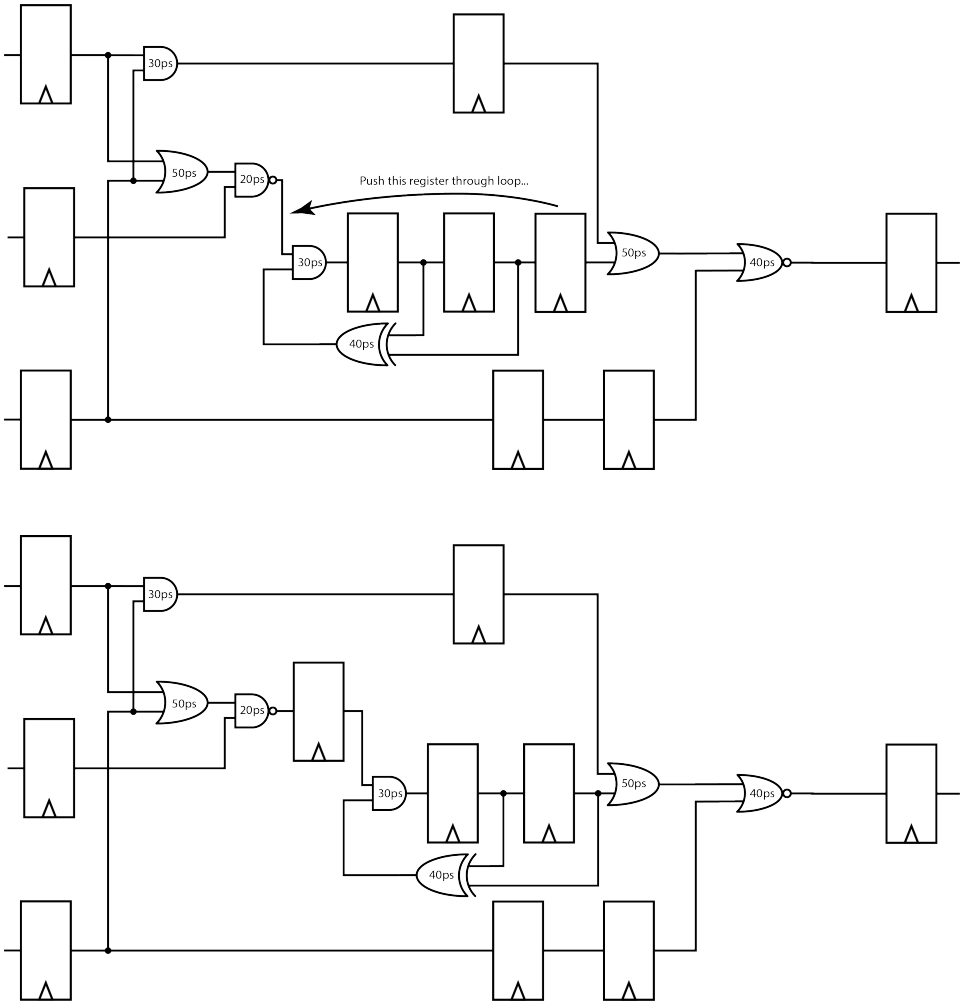



To break this path, we will have to keep pushing the register through the design.



Since there is a loop involved, we will want to keep this loop intact during the retiming, so we treat this loop as a fixed unit and push the register directly to the input to the loop.



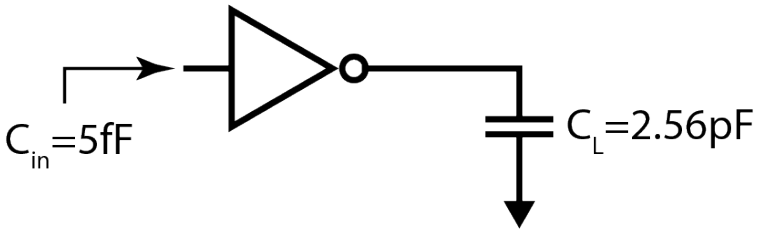

The final critical path of this network is thus 90 ps, and so the minimum clock period we can tolerate is,

$$T_{min} \geq t_{\text{clk-q}} + t_{p,\text{crit}} + t_{\text{setup}} = 10\,\text{ps} + 90\,\text{ps} + 20\,\text{ps} = 120\,\text{ps}$$

which leads to a maximum clock frequency of 8.3 GHz.

## Problem 2: Buffer Insertion

After finishing the place-and-route process for your new ASIC design, you notice that due to a mistake in your constraint setup one of your signals is attempting to drive a very large gate directly using a minimum-sized inverter. After doing a bit of inspection of your design, you draw up the following schematic for the loading on the inverter output.




Find the optimal number of buffer stages and buffer sizes to insert between the minimum-sized inverter and the load to minimize the delay on this connection. In this technology, $\gamma = 1.5$. *You must preserve the logical function at the output as this wire carries a signal.*

Solution:

First, find the optimal fanout. Since we are given $\gamma = 1.5$, we can solve the transcendental equation for optimal fanout,

$$f = e^{(1+1.5/f)} \rightarrow f \approx 3.97$$

Which we can approximate as $f = 4$. This will be our target for the fanout per buffer stage. Following the steps on Slide 24 Lecture 12,

$$f = \sqrt[N]{F} = \sqrt[N]{\frac{2.56\,\text{pF}}{5\,\text{fF}}} = \sqrt[N]{512} = 4$$

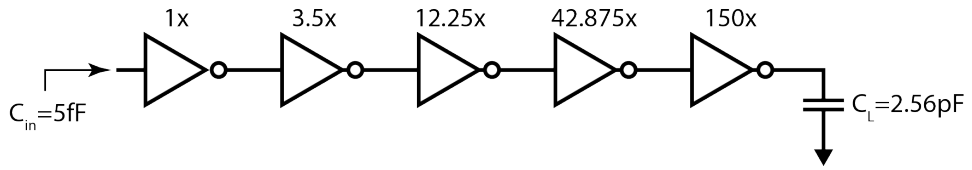
Solving for N,

$$\begin{aligned} 4 &= \sqrt[N]{512} \\ 4^N &= 512 \\ N &= \log_4 512 = 4.5 \end{aligned}$$

We can't have half an inverter, so we will need to either round up or round down. Since we want to preserve the logical function of the original line (driven by 1 inverter), we need an odd number of drivers, so we will pick $N = 5$. This means we will need to redistribute the fanout amongst the stages again, so the final fanout per stage will be

$$f = \sqrt[5]{512} \approx 3.5$$

which results in this final setup.



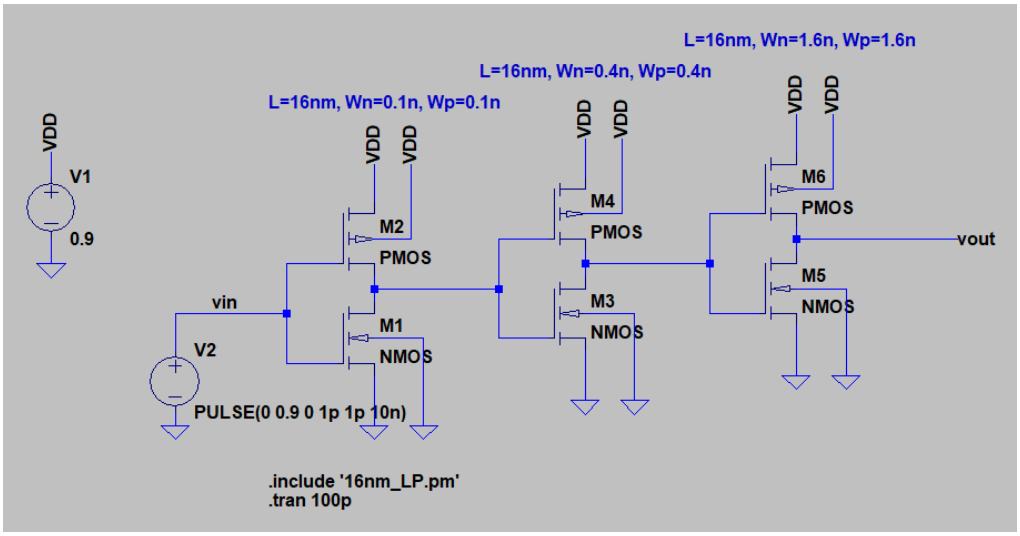

## Problem 3: FO4 Delay

Using the SPICE model library from previous homeworks, set up a schematic testbench to find the FO4 delay for this technology, with $W_N = W_P = 100\,\text{nm}$. Remember that this setup means an

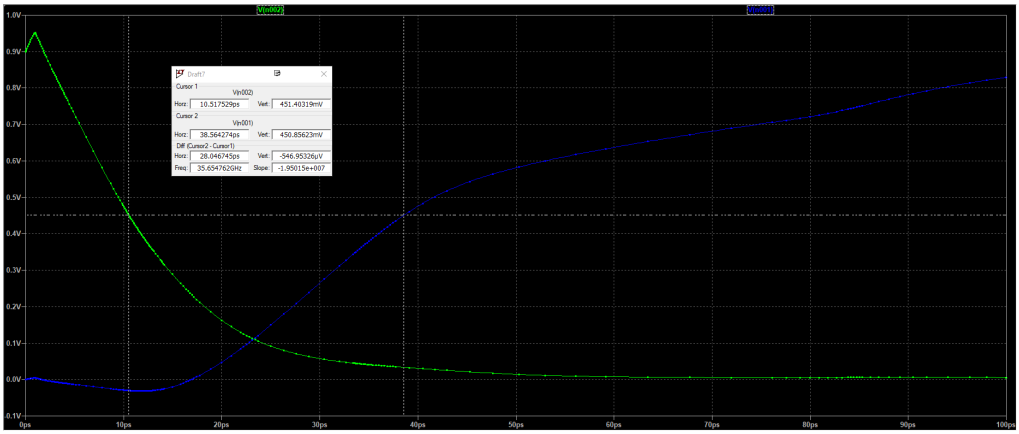
inverter driving one 4x bigger, with that inverter driving another inverter that is 4x bigger that it and measuring the delay across the inverter in the middle. Repeat with $W_N = W_P = 500\,\text{nm}$. Is the delay the same? Is this what you expect?

**Solution:**

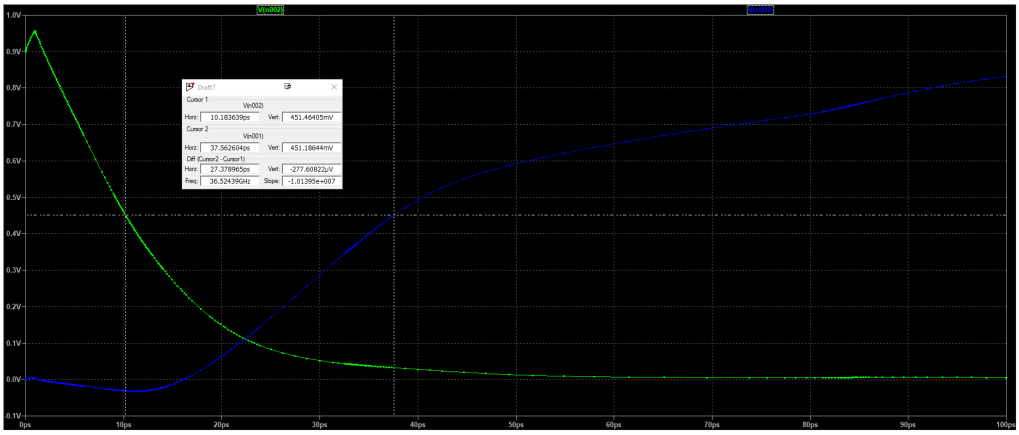
Schematic for an inverter, driven by an inverter 4x smaller than itself, and driving an inverter 4x larger than itself:




Finding the delay of the middle inverter, we get roughly 28 ps.



Repeating for $W = 0.5\,\mu\text{m}$,



We see that the FO4 delay is roughly the same, though there are some small differences. This is expected as the FO4 delay should be roughly constant as it is defined in relation to a ratio of sizes. The small differences can be attributed to the source/drain parasistic capacitances that do not scale linearly with device width.

## Problem 4: Gate Delay

Following the same steps as the gate delay calculation in the lecture, find the gate delays for the following gates. In this technology, $R_P = 2R_N$. (*HW Note: As an alternative to a short explanation at the end of your derivation, you may explain each step of your derivation as you present your work.*)

(a) 4-input NAND

Solution:

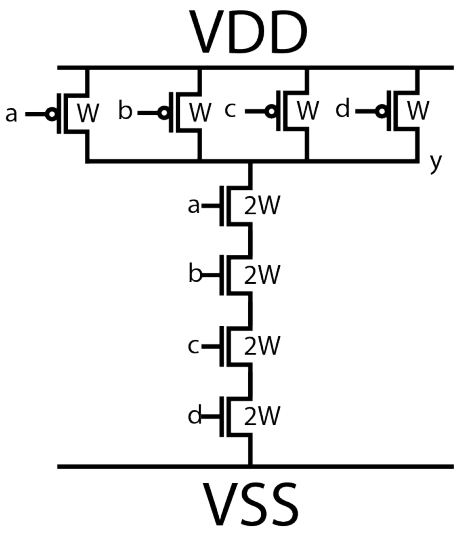



Figure 1: 4-NAND CMOS Schematic

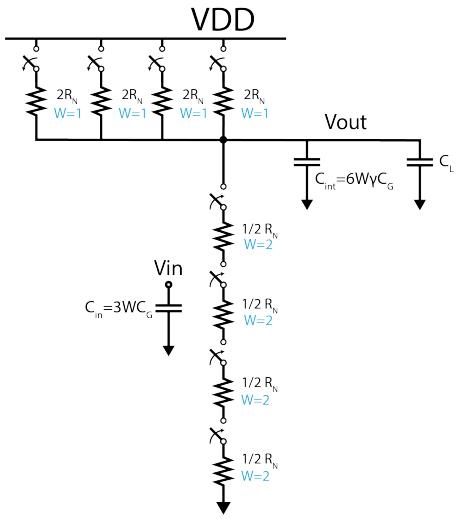



Figure 2: 4-NAND Gate Delay Schematic

$$C_{int} = 4\gamma C_G + 2\gamma C_G = 6\gamma C_G$$

$$\begin{aligned} t_p &= 0.69\left(\frac{2R_N}{W}\right)(6\gamma C_G + C_L) \\ &= 0.69\left(\frac{2R_N}{W}\right)(3\gamma W C_G)\left(2 + \frac{C_L}{3W\gamma C_G}\right) \\ &= t_{p0}\left(4 + \frac{2f}{\gamma}\right) \end{aligned}$$

(b) 3-input NOR

Solution:

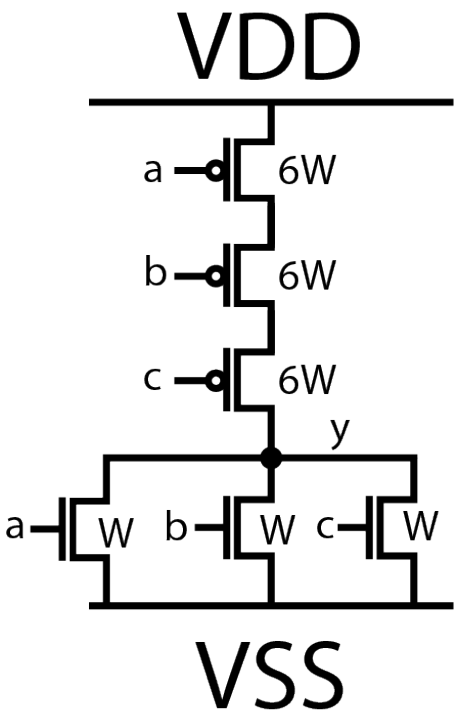



Figure 3: 3-NOR CMOS Schematic

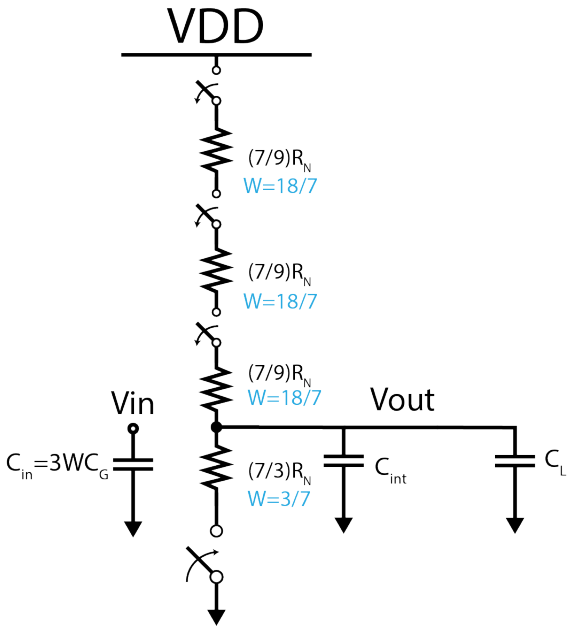



Figure 4: 3-NOR Gate Delay Schematic

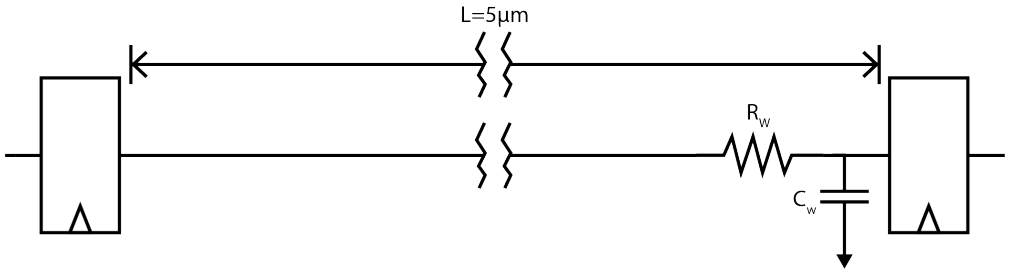
$$C_{int} = (18/7)\gamma C_G + (9/7)\gamma C_G = (27/7)\gamma C_G$$

$$\begin{aligned} t_p &= 0.69\left(\frac{7R_N}{3W}\right)\left(\frac{27\gamma C_G}{7} + C_L\right) \\ &= 0.69\left(\frac{7R_N}{3W}\right)(3\gamma W C_G)\left(\frac{9}{7} + \frac{C_L}{3W\gamma C_G}\right) \\ &= t_{p0}\left(9 + \frac{7f}{\gamma}\right) \end{aligned}$$

How would your answers change if $R_P = R_N$?

## Problem 5: Wire Rebuffering

After working on your design for hours and finally putting it through the place-and-router, you're greeted with the sight you least want to see: your design failed to close timing. Heaving a heavy sigh, you open up the timing report to find that two registers on your design were placed so far away that the signal from one is not able to reach the other within a clock cycle. From the report, you are able to gather the following schematic of the connection.




In order to correct this problem, you will need to add buffers along the long wire ($r_W = 0.002\,\Omega/\mu\text{m}$, $c_W = 5 \times 10^{-3}\,\text{F/m}$) between the launching and receiving flip-flops. Find the minimum number of buffers you can add to make this path meet timing at a clock frequency of 1 GHz. You may use any number of inverters for buffering as any inversion in the line can be made up for elsewhere. You may assume the flip-flop drives its output with the same strength and output capacitance as a minimum-sized inverter ($R_{dr} = 500\,\Omega$, $C_{int} = 2\,\text{fF}$).

Solution:

First, an apology for this question. The wire properties were not properly vetted before releasing this question, which led to significant confusion as to how to solve this problem. I will present the intended solution with updated parameters.

If we instead consider the wire to be 500 µm with a resistivity of $50\,\Omega/\mu\text{m}$ and a capacitivity of 1 nF/m, we can approach this problem by trying to subdivide the wire into smaller segments.

The equation for the delay along a long wire driving a load is as follows from the lecture.

$$t_p = 0.69R_{dr}C_{int} + 0.69R_{dr}C_w + 0.69R_{dr}C_{fan} + 0.69R_wC_{fan} + 0.38R_wC_w \quad (1)$$
$$= 0.69R_{dr}C_{int} + 0.69R_{dr}(c_wL) + 0.69R_{dr}C_{fan} + 0.69(r_wL)C_{fan} + 0.38(r_wc_wL^2) \quad (2)$$

Evaluating this equation for $L = 500\,\mu m$ with the constants given, we find that this connection has a delay of 4.96 ns, which will not be enough to make the 1 GHz clock period, even when assuming the registers are ideal. By adding a buffer in the middle of the wire, we can now treat this problem two long wire drivers, each driving a wire half as long.

$$t_p = 2 * \left(0.69R_{dr}C_{int} + 0.69R_{dr}\left(c_w\left(\frac{L}{2}\right)\right) + 0.69R_{dr}C_{fan} + 0.69\left(r_w\left(\frac{L}{2}\right)\right)C_{fan} + 0.38\left(r_wc_w\left(\frac{L}{2}\right)^2\right)\right)$$

This significantly decreases the delay along this entire line as the minimum-sized drivers are now sharing the large wire by dividing it. Evaluating the delay, we see that this new network now has a delay of 2.58 ns. At fist glance, we would estimate this strategy would divide our delay in half, which we would expect to be 2.48 ps. However, the actual delay is a little bit higher because the new driver itself introduces some intrinsic delay which does not scale with L. Continuing this pattern, we find that at $N = 6$, the delay of the connection finally goes below 1 ns and so we will need 6 buffers along this line to meet the time constraint.

## Problem 6: Delays in a Physical Circuit

While performing a design review for your fellow layout engineers, you notice a long route between two identical inverters that may present a problem, illustrated below with dimensions annotated.

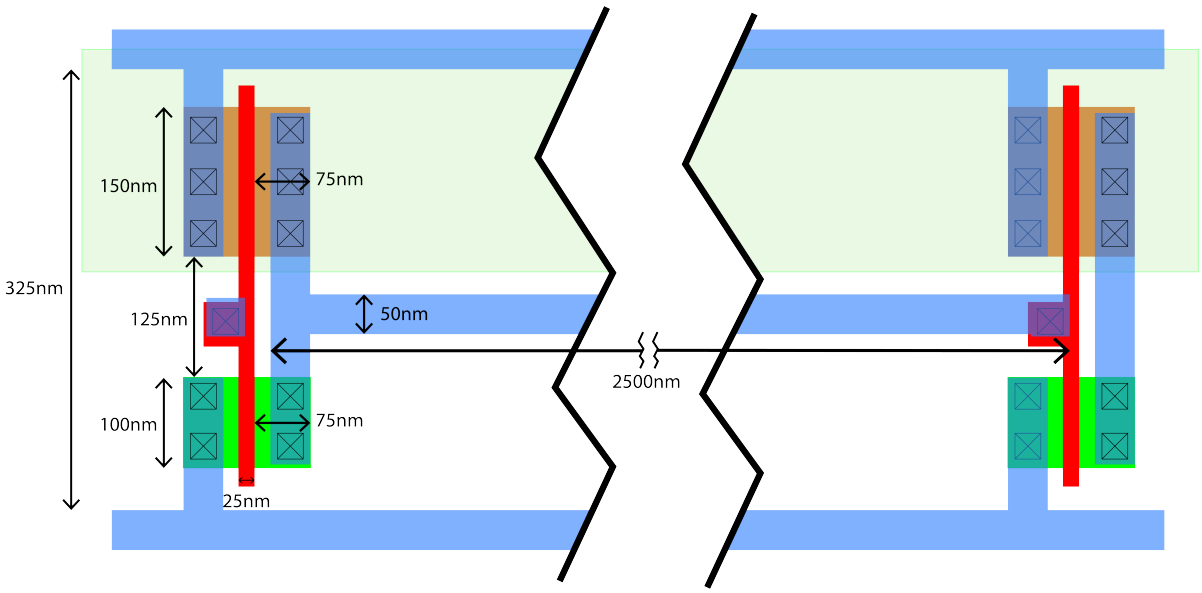



Figure 5: Layout With Annotated Dimensions

This technology has the following process parameters

| Parameter | Description | Value |
|---|---|---|
| $C_j$ | Bottom junction capacitance per unit area | $0.2\,\mathrm{F/m^2}$ |
| $C_{jsw}$ | Sidewall perimeter capacitance per unit length | $1.6\,\mathrm{F/nm}$ |
| $C_{ox}$ | Gate-bulk oxide capacitance per unit area | $0.016\,\mathrm{F/m^2}$ |
| $R_{\square,\mathrm{M1}}$ | Metal 1 sheet resistance | $4\,\mathrm{m\Omega/\square}$ |
| $R_{\square,\mathrm{pdiff}}$ | p+ diffusion sheet resistance | $0.002\,\mathrm{m\Omega/\square}$ |
| $R_{\square,\mathrm{ndiff}}$ | n+ diffusion sheet resistance | $0.001\,\mathrm{m\Omega/\square}$ |
| $R_{\square,\mathrm{Poly-Si}}$ | Polysilicon sheet resistance | $0.00001\,\mathrm{m\Omega/\square}$ |
| $c_w$ | Wire parasitic capacitance per unit area | $4\,\mathrm{mF/m^2}$ |
| $R_{DSN,min}$ | Minimum-size NMOS source-drain resistance | $700\,\Omega$ at $W = 100\,\mathrm{nm}$ |
| $R_{DSP,min}$ | Minimum-size PMOS source-drain resistance | $1050\,\Omega$ at $W = 100\,\mathrm{nm}$ |

Table 1: Process Parameters

Find the propagation delay of the first inverter given these parameters and the layout dimensions. You may assume that there is negligible parasitic contribution from the vias, source/drain metal contacts, supply rails, and polysilicon **where it does not overlap the diffusion regions**. You must take into account the parasitic capacitances of the poly-diffusion overlap areas, however.

Solution:

Again, there was some confusion about the units in this problem, in particular the sidewall capacitance of the source/drain regions. The number I provided here is far too high as the junction capacitance is usually in the femtofarads, so the answer you would get from this problem is unrealistic. I will present the solution with a more realistic junction capacitance.

To begin, we need to find all the necessary parameters to calculate long wire delay.

- $R_{dr}$

  The driving resistance of this gate is found by dividing $R_{DSN,min}$ or $R_{DSP,min}$ by the width of the transistor, normalized to a minimum width. The NMOS in this inverter is already at minimum width, so immediately we know $R_{DSN} = 700\,\Omega$. The PMOS is 1.5 times larger than the minimum, so we can calculate

  $$R_{DSP} = R_{DSP,min}/1.5 = 700\,\Omega$$

  so the gate is symmetric and $R_{dr} = 700\,\Omega$.

- $R_w$

  The resistance of the wire is found by using the sheet resistance value given to find the series resistance of the wire. The long wire is 50 nm wide but 2500 nm long, which means the area of the wire can be divided in to 50 50 nmx50 nm squares, leading to a resistance of $200\,\Omega$. We can consider the wire routing from the NMOS and PMOS drains to be of neglible contribution to resistance, so we will only consider the long wire.

- $C_w$

To find the wire capacitance, find the area of the wire and multiply by the given capacitance per unit area.

$$C_w = (2500\,\text{nm})(50\,\text{nm})(4\,\text{mF/m}^2) = 500\,\text{aF}$$

Note that in a real layout, there will be some significant fringing capacitance from the sides of the wire which will likely be on the same order as the parasitic capacitance itself. We can model that in this problem by doubling the wire capacitance, so $C_w = 1\,\text{fF}$. This part was not mentioned explicitly in this problem or the discussion, but in future problems the fringing capacitance term will be given in the parameter table.

- $C_{in}$

  For the input capacitance of the receiving inverter, we find the gate capacitance of both the PMOS and NMOS devices by finding the gate area and multiplying by the oxide capacitance. For the PMOS,

  $$C_P = (150\,\text{nm})(25\,\text{nm})(0.016\,\text{F/m}^2) = 60\,\text{aF}$$

  For the NMOS,

  $$C_P = (100\,\text{nm})(25\,\text{nm})(0.016\,\text{F/m}^2) = 40\,\text{aF}$$

  so together, $C_{in} = 100\,\text{aF}$.

- $C_{int}$

  This is the term with the error in the parameter table. For this solution, we will be assuming the (more realistic) value of $0.08\,\text{nF/m}$. The capacitance seen from the PMOS drain is the perimeter of the drain diffusion region (without the channel-facing side) multiplied by the sidewall capacitance per unit length.

  $$C_{jsw,P} = (150\,\text{nm} + 2(75\,\text{nm}))(0.08\,\text{F/m}) = 24\,\text{aF}$$

  The NMOS capacitance is similarly found as

  $$C_{jsw,N} = (100\,\text{nm} + 2(75\,\text{nm}))(0.08\,\text{F/m}) = 20\,\text{aF}$$

  We now factor in the bottom junction capacitance for each transistor.

  $$C_{jP} = (150\,\text{nm})(75\,\text{nm})(0.2\,\text{F/m}^2) = 2.25\,\text{fF}$$

  $$C_{jN} = (100\,\text{nm})(75\,\text{nm})(0.2\,\text{F/m}^2) = 1.5\,\text{fF}$$

  Together the capacitance at the output is $3.794\,\text{fF}$.

Using these terms, we can plug them into our equation for long wire delay,

$$t_p = 0.69R_{dr}C_{int} + 0.69R_{dr}C_w + 0.69R_{dr}C_{in} + 0.69R_wC_{in} + 0.38R_wC_w$$

Evaluating, we arrive at $t_p = 2.52\,\text{ps}$.