Solution to Mid Semester Examination
EE 671: VLSI Design: Autumn Semester 2016

Q–1 In a given process, the ratio of p channel transistor width to the n channel transistor width in an inverter should be 4 to obtain equal rise and fall times. The parasitic delay of the inverter is 2.3.

a) The ideal stage ratio ρ is a solution to the equation $p_{inv} + \rho(1 - \ln \rho) = 0$. Evaluate the value of ρ using the Newton Raphson iterative technique, starting with a guess value of 4. Values for ρ, $f(\rho)$ and $f'(\rho)$ should be reported for each iteration, till you reach a convergence to 3 decimal places. (The reference section gives a brief description of Newton Raphson method.)

Soln. 1-a)

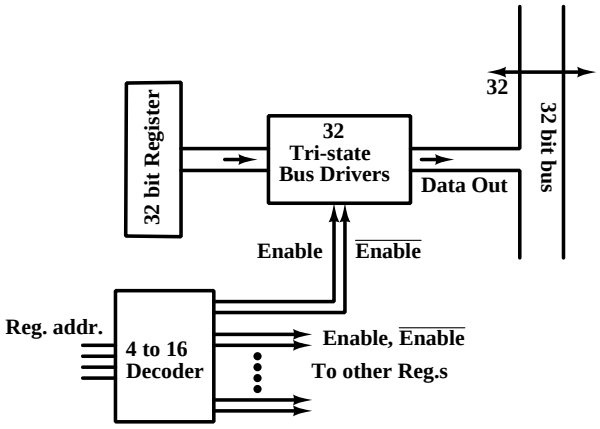
$$\begin{aligned} f(\rho) &= \rho(1 - \ln \rho) + p_{inv} \\ f'(\rho) &= 1 - \ln \rho + \rho(-\frac{1}{\rho}) = -\ln \rho \end{aligned}$$

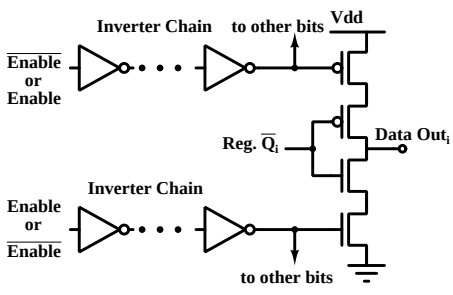
ρ	$f(\rho)$	$f'(\rho)$	next ρ
4.0	0.7548	-1.3863	4.5445
4.5445	-0.0355	-1.5139	4.5211
4.5211	-6.0544×10^{-05}	-1.5087	4.5210
4.5210	-1.7809×10^{-10}	-1.5087	4.5210
4.5210	0	-1.5087	4.5210

– **[3]**

b) The figure on the left below shows a scheme to couple the output of a selected 32 bit register to a bus. Outputs from the decoder (*Enable* / $\overline{Enable}$) drive inverter chains, which in turn drive 32 tri-state output stages (one for each bit). The figure on the right shows inverter chains driving output stages. (Only one output stage is shown, the remaining 31 are indicated by the arrow).

Each tri-state output stage drives a line of the 32 bit bus. The capacitive load presented by the bus line is equivalent to the input capacitance of 40 minimum inverters.





Find the number of inverters we should insert between the p channel transistor

gate and the decoder output to minimize the total delay. (Notice that we have the option of selecting either the *Enable* or the $\overline{Enable}$ output of the decoder depending on whether an even or an odd number of inverters are required).
Assume that *Enable*/$\overline{Enable}$ outputs of the decoder can drive two minimum inverters each.

Soln. 1-b) The tri-state output stage has two n channel transistors in series. Their widths should be 2 each (in units of minimum sized transistors). Two p channel transistors are also in series. Since γ is 4, the p channel transistors should have widths of 8 each. An inverter will have its n channel transistor with width of 1 and the p channel transistor with a width of 4, for a total capacitive load of 5 units. Thus the logical effort for driving the p channel transistor of the output stage is 8/5, whereas for driving the n channel transistor, it is 2/5.
Consider the path from the decoder output to the bus output through the p channel transistor. To compute the path effort, $G = 1 \times 1 \times \cdots 8/5 = 1.6$
$H = 40/2 = 20$
$B = 32$
Therefore the path effort is $F = GBH = 1.6 \times 32 \times 20 = 1024$
The optimum number of stages is
$N = \ln(F)/\ln(\rho) = \ln(1024)/\ln(4.5210) = 4.5942$,
So we require 5 stages in the path. One of these is the output stage itself. Therefore we need to insert 4 inverters. This will be non-inverting, so we should choose the $\overline{Enable}$ output of the decoder to drive this path. – **[3]**

c) Find the scale factor for each of the inverters inserted in the inverter chain. Compute the delay for the path from decoder output to the bus wire in units of τ.

Soln. 1-c) We have total path effort $F = 1028$ and 5 stages. Therefore each stage will have an effort $\hat{f} = 1028^{1/5} = 4$. Let us start with the output stage. Its stage effort must be 4. Therefore,

$$gbh = 8/5 \times 1 \times 40/C_{in} = 4, \text{ so } C_{in} = \frac{8}{5} \times 40/4 = 16$$

Thus this stage should have an input capacitance of 16 inverters. Therefore its scale factor is $16/1.6 = 10$.

We next take up the last of the four inverters, Inv_4. For this stage, $g = 1, b = 32, h = 16/C_{in}$. So we must have

$$32 \times 16/C_{in} = 4, \text{ which gives } \quad C_{in} = 32 \times 16/4 = 128$$

.
The third inverter has $g = 1, b = 1$ and it has to drive a load of 128 inverters. Since the stage effort should be 4, its input capacitance is given by

$$128/C_{in} = 4 \text{ so } \quad C_{in} = 32$$

The second inverter also has $g = 1, b = 1$ and drives a load of 32 inverters. So its input capacitance is $32/4 = 8$.

The first inverter drives a load of 8 inverters and has $g = 1, b = 1$. Therefore its input capacitance is $8/4 = 2$. This agrees with the driving capability of $\overline{Enable}$ output of the decoder.

So the scale factors of the various stages are:

Stage →	1	2	3	4	5
Type	Inv.	Inv.	Inv.	Inv.	Tri-stage
Scale	2	8	32	128	10

The toal delay of this path is $5 \times 4+$ parasitic delay.
The parasitic delay of each inverter is 2.3. The unit tri-state driver has an n channel transistor of geometry 2 and a p channel transistor of geometry 8 connected to the output. This is twice the parasitic capacitance of a minimum inverter (which has $1 + 4 =$) 5. So we can estimate the parasitic delay of the tri-state driver to be twice that of the inverter. Hence we can take the parasitic delay of the tri-state driver to be 4.6. Then the total delay is $20 + 4 \times 2.3 + 4.6 = 33.8$. – [4]

d) How many inverters do we need to insert to drive the n channel transistor gates? Depending on which of *Enable* or $\overline{Enable}$ was chosen for the part above, the other output of the decoder should be used for this chain. Adjust the number of inverters for this and compute the total delay for this path.
Compute the scale factors for all inverters in this chain.
There is no need to equalize delays through the two chains.

Soln. 1-d) To compute the total path effort, the logical effort of the tri-state buffer is to be computed for the n channel transistors. The relative size of the n channel transistor is 2, while the total capacitive load presented by a minimum inverter is $(1+4 = 5)$. Therefore, the logical effort for the n channel transistors is $2/5$. Then, we have $G = 1 \times 1 \times \cdots \times 2/5 = 0.4$. As before, $B = 32$ and $H = 40/2 = 20$. So the path effort is $F = GBH = 0.4 \times 32 \times 20 = 256$

Since the equality of rise time and fall time must be maintained, the scale of the tri-state driver is already fixed by the design of p channel drivers. So the capacitance presented by the tri-state driver to the inverter chain is $10 \times 2 = 20$ minimum transistor widths. In units of inverter capacitances, this is equivalent to 4 inverters. So the stage effort for the tri-state driver is $0.4 \times 1 \times 40/4 = 4$. The path effort from the input to the gate of the n channel transistor in the tri-state buffer is $256/4 = 64$. (This could be calculated independently. The total load on the final inverter inclusive of branching is $32 \times 4 = 128$. Since the input capacitance is 2, $BH = 128/2 = 64$. The logical effort of all inverters is 1, so the path effort is $F = GBH = 64$, excluding the final tri-state buffer.)

The number of inverters to be inserted is given by $\ln 64 / \ln 4.521 = 2.7565$. So we would like to have 3 inverters. However, that would mean that the inverter chain will be inverting and we must use $\overline{Enable}$ as the input. This option is not available to us. We have to use *Enable* as the input and therefore, must have an even number of inverters in the driver chain. So we choose 4 inverters in the chain. (A choice

of 2 inverters is actually closer to 2.76, so we should also consider using 2 inverters).

The optimum stage ratio considering only the 4 inverters is $\hat{f} = 64^{1/4} = 2.8284$. Since the total load on the fourth inverter is $32 \times 4 = 128$ and g and b values for all other stages are 1, we get the scale factors for all stages by dividing 128 repeatedly by $\hat{f}$. So we get the scale factors as 45.26, 16, 5.66 and 2 respectively. The total delay is then

$$4\hat{f} + 4 + 4p_{inv} + p_{tristate} = 11.3136 + 4 + 9.2 + 4.6 = 29.11\tau$$

Additional discussion

Notice that the parasitic delay accounts for about half the buffer delay, inspite of the heavy loading in this problem. It is interesting to check what would happen if chose two stages of inverters instead of four.

In that case, $\hat{f} = 64^{1/2} = 8$ and the scale factors will be 2 and 16 for the two inverters. The total delay in this case will be

$$2\hat{f} + 2p_{inv} + f_{tristate} + p_{tristate} = 16 + 4.6 + 4 + 4.6 = 29.2\tau$$

This is practically the same as the four inverter case. This happens because 2 and 4 are about equidistant from the optimal choice of 3 inverters.

What if we did not respect the p to n ratio in the tri-state buffer?
In that case, the path effort of 256 can be spread equally over all 5 stages. So $\hat{f} = 256^{1/5} = 3.031$. The stage effort for the tri-state buffer is

$$\hat{f} = 3.031 = gbh = 0.4 \times 1 \times 40/C_{in} = 16/C_{in}$$

This gives $C_{in} = 16/3.031 = 5.28$ An input capacitance of 5.28 inverters corresponds to $5.28 \times 5 = 26.4$ minimum transistor widths. So the scale factor is 13.2.

The load on the fourth inverter inclusive of branching is $32 \times 5.28 = 168.9$. Therefore the input capacitance of the fourth inverter is $168.9/3.031 = 55.72$.

Successively dividing by 3.031, we get the scale factors of $55.72/3.03 = 18.38$ for the third, and $18.38/3.03 = 6.062$ for the second, and finally, $6.062/3.031 = 2$ for the first inverter.

The total delay is

$$5 \times 3.031 + 4 \times 2.3 + 4.6 = 28.96$$

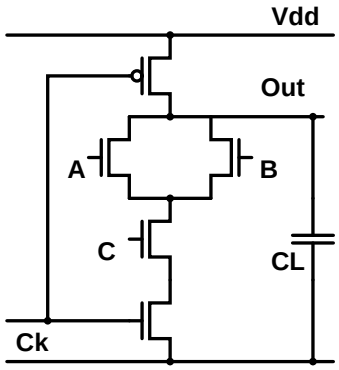
This is not very different from the designs which do maintain equal rise and fall times. So those must be preferred over this design. – [4]

– **[Q1: 3 + 3 + 4 + 4 = 14 marks]**

Q–2 Why does a CMOS dynamic logic gate malfunction if we do not use multiple clocks? How is this problem solved using 4 phase dynamic logic? What is the restriction for driving different types of gates in 4 phase logic?

How does Zipper logic manage to solve this problem without needing multiples phases of the clock?

Soln. 2) Let us consider a basic CMOS dynamic gate inplementing $\overline{(A+B).C}$.

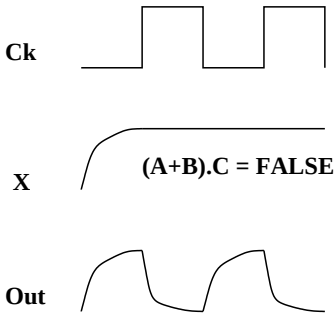


When the clock is low, the pMOS transistor is on and the bottom nMOS is off. The output is 'pre-charged' to '1' unconditionally.

When the clock goes high, the pMOS turns off and the bottom nMOS comes on.
The circuit then conditionally discharges the output node, if (A+B).C is TRUE.

This implements the function $\overline{(A+B).C}$.

However this circuit can malfunction when several CMOS dynamic gates are cascaded. Consider the case when the above circuit is followed by a dynamic inverter.

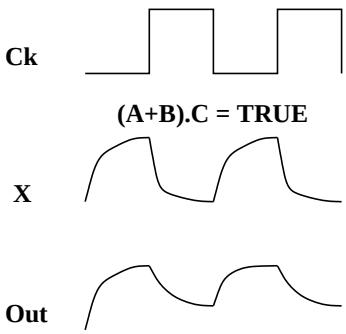


When (A+B).C is FALSE, There is no problem.

X pre-charges to '1' and remains at '1'..

Therefore the inverter sees the correct logic value at its input all the time and discharges the output to '0', which is the expected output.

However, When (A+B).C is TRUE, there is a problem.



The correct value for X is now '0'. After the pre-charge cycle, X takes some time to discharge to '0'. So for some time after pre-charge, its output is held at the wrong value of '1'.

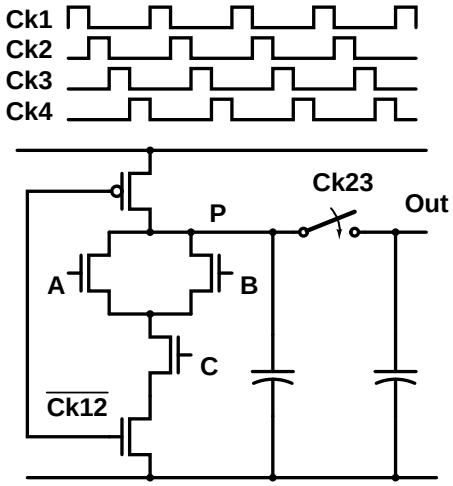
During this time, the nMOS of the inverter is 'on' and charge placed on the output leaks away. This can take the output to the wrong value of '0'. There is no way for the output to go to '1' when X reaches its correct value of '0'.

Thus the simple dynamic CMOS circuit can malfunction on cascading for certain data conditions.

This problem can be soved by using a clock with multiple phases, so that the transiently wrong output of one stage is not fed to the other. We use different phases for pre-charge,

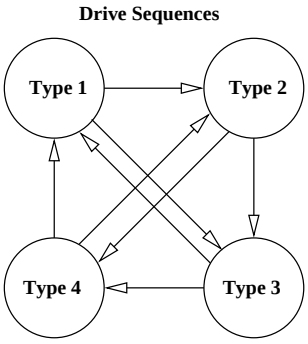
evaluation and for holding a valid output. Now we can sample the previous stage only in the clock phase when evaluation is complete and it is valid.

For the circuit shown below, we use a 4 phase clock. Combined clock signals of the type Ck_{mn} are generated as required. Ck_{mn} is high during the m and n phases of the clock.



1. In phase 1, node P is pre-charged.
2. In phase 2, P as well as the output are pre-charged.
3. In phase 3, The gate evaluates.
4. In phases 4 and 1, the output is isolated from the driver and remains valid.

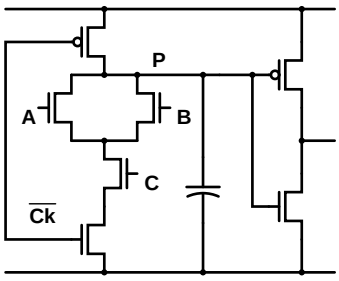
This is called a type 3 gate. It evaluates in phase 3 and is valid in phases 4 and 1. Similarly, we can have type 4, type 1 and type 2 gates.



Since the output of type 3 gate is valid in phase 4 and phase 1, it can drive gates of type 4 and type 1.
Similarly, a type 1 gate can feed gates of types 2 and 3; type 2 gates can drives types 3 and 4; and type 4 gates can drive types 1 and 2.

This ensures that gates receive inputs only when these are valid.

The CMOS dynamic logic gate malfunctions because the output remains at the pre-charge value of '1' for some time, before acquiring it valid value of '0'. We can also solve this problem by putting a static inverter after the CMOS dynamic gate. When the logic gate pre-charges to '1', the output of the inverter goes to '0'.

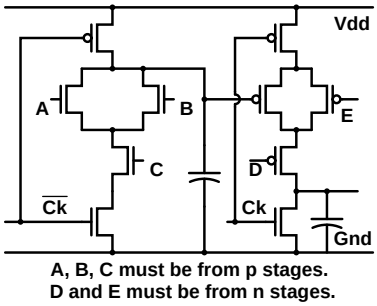


The output of the inverter goes to '0' on pre-charge and may remain at the wrong value of '0' for some time before acquiring its right value of '1'. However this wrong value does not turn on the nMOS input of the next stage and therefore does not lead to a malfunction.

This circuit is non-inverting because of the addition of an inverter. So all logic functions cannot be implemented using it.

We can implement arbitrary logic functions using zipper logic. Just as a 'transiently wrong' value of '0' does not cause a problem for nMOS inputs, a 'transiently wrong'

value of '1' is safe for pMOS inputs. Therefore we can alternate n and p type CMOS dynamic logic without using an inverter.



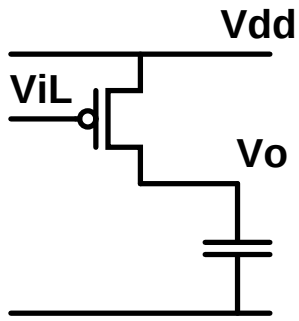
The n stage is pre-charged high, but it drives a p stage. A high pre-charged stage will keep the p evaluation stage off, which will not cause any malfunction. The p stage will be pre-discharged to 'low', which is safe for driving n stages.

– **[7 marks]**

Q–3 Consider a CMOS inverter in which n and p channel transistors have been sized to give equal rise and fall times. Derive an expression in terms of p channel transistor parameters for charging the output capacitance C_L from 0V to 3V with $V_{DD} = 3.3$V. Assume the input voltage to be 0.5V, which is below V_{Tn}. You can assume perfect saturation.

If the value of $K_p \equiv \mu C_{ox} W/L$ is 100μA/V^2, $V_{DD} = 3.3$V, $V_{Tp} = 0.7$V and the load capacitance is 0.1 pF, find the charge time from 0 to 3V. Find the value of the equivalent resistor which will charge the load capacitance in the same amount of time from 0 to 3V.

Soln. 3) Because the input voltage is $< V_{Tn}$, the nMOS is off and need not be considered.



$$I_{dp} = C_L \frac{dV_o}{dt}$$

so,

$$\frac{dt}{C_L} = \frac{dV_o}{I_{dp}}$$

Integrating both sides, we get

$$\frac{\tau_{rise}}{C_L} = \int_0^3 \frac{dV_o}{I_{dp}}$$

The pMOS is saturated till the output reaches $0.5+V_{Tp}$. Between $0.5+V_{Tp}$ to 3.0 V, it is in linear regime. The magnitude of V_{GS} for the p channel transistor is $3.3 - 0.5 = 2.8$V.

$$\begin{aligned}\frac{\tau_{rise}}{C_L} &= \int_0^{0.5+V_{Tp}} \frac{dV_o}{\frac{K_p}{2}(2.8 - V_{Tp})^2} \\ &\quad + \int_{0.5+V_{Tp}}^{3} \frac{dV_o}{K_p\left[(2.8 - V_{Tp})(3.3 - V_o) - \frac{1}{2}(3.3 - V_o)^2\right]}\end{aligned}$$

We define $V_x \equiv 3.3 - V_o$. Then $\mathrm{d}V_o = -\mathrm{d}V_x$. As V_o goes from $0.5 + V_{Tp}$ to 3.0V, V_x will go from $2.8 - V_{Tp}$ to 0.3V.

$$\begin{aligned}\frac{K_p \tau_{rise}}{2C_L} &= \frac{0.5 + V_{Tp}}{(2.8 - V_{Tp})^2} - \int_{2.8-V_{Tp}}^{0.3} \frac{\mathrm{d}V_x}{2V_x(2.8 - V_{Tp}) - V_x^2} \\ &= \frac{0.5 + V_{Tp}}{(2.8 - V_{Tp})^2} + \int_{0.3}^{2.8-V_{Tp}} \frac{\mathrm{d}V_x}{V_x\,(5.6 - 2V_{Tp} - V_x)}\end{aligned}$$

The integration can be carried out using partial fractions.

$$\int_{0.3}^{2.8-V_{Tp}} \frac{\mathrm{d}V_x}{V_x\,(5.6-2V_{Tp}-V_x)} = \frac{1}{5.6-2V_{Tp}} \int_{0.3}^{2.8-V_{Tp}} \left(\frac{1}{5.6-2V_{Tp}-V_x}+\frac{1}{V_x}\right)\mathrm{d}V_o$$

$$= \frac{1}{5.6-2V_{Tp}} \left[\ln \frac{V_x}{5.6-2V_{Tp}-V_x}\right]_{0.3}^{2.8-V_{Tp}}$$

$$= \frac{1}{5.6-2V_{Tp}} \ln \frac{5.3-2V_{Tp}}{0.3}$$

So

$$\frac{K_p \tau_{rise}}{2C_L} = \frac{0.5+V_{Tp}}{(2.8-V_{Tp})^2} + \frac{1}{5.6-2V_{Tp}} \ln \frac{5.3-2V_{Tp}}{0.3}$$

Thus

$$\tau_{rise} = \frac{2C_L}{K_p}\left(\frac{0.5+V_{Tp}}{(2.8-V_{Tp})^2} + \frac{1}{5.6-2V_{Tp}} \ln \frac{5.3-2V_{Tp}}{0.3}\right)$$

This is the expression for τ_{rise} in terms of device parameters.

If the value of $K_p \equiv \mu C_{ox} W/L$ is given to be $100\mu\mathrm{A/V}^2$, $V_{DD} = 3.3\mathrm{V}$, $V_{Tp} = 0.7\mathrm{V}$ and the load capacitance = 0.1 pF, The above expression evaluates to

$$\frac{2\times 10^{-13}}{10^{-4}}\left(\frac{1.2}{2.1^2}+\frac{1}{4.2}\ln\frac{3.9}{0.3}\right) = 1.6861\mathrm{ns}$$

If a resistor charges the same capacitor from 0V to 3V in the same time, we should have

$$3 = 3.3(1-e^{-\tau_{rise}/RC_L}) \quad \text{So} \quad 1-e^{-\tau_{rise}/RC_L} = \frac{3}{3.3}$$

Which gives

$$e^{-\tau_{rise}/RC_L} = 1-\frac{3}{3.3} = \frac{.3}{3.3} = \frac{1}{11} \qquad \text{So} \quad e^{\tau_{rise}/RC_L} = 11$$

Thus,

$$\frac{\tau_{rise}}{RC_L} = \ln(11) \quad \text{and so,} \quad R = \frac{\tau_{rise}}{\ln(11)\times C_L}$$

Since $\tau_{rise} = 1.6861\mathrm{ns}$ and $C_L = 0.1\mathrm{pF}$, we can evaluate R as

$$R = \frac{1.6861\times 10^{-9}}{2.3979\times 10^{-13}} = 7.0399\times 10^3$$

So the equivalent resistor is 7.04KΩ. – **[6 marks]**

Q–4 What is the value of Metal to Semiconductor work function ϕ_{MS} if we are using Aluminum as the metal with p type silicon doped to $10^{17}/\mathrm{cm}^3$ as the substrate? The Fermi level of silicon is 50 meV above the conduction band of silicon. The band gap of silicon is 1.1 eV and the intrinsic carrier density in silicon is $1.5\times 10^{10}/\mathrm{cm}^3$. Take the value of KT/q to be 26 meV.

Soln. 4) The Fermi level in silicon will be $(KT/q)\ln(N_A/n_i)$ below the intrinsic level. Thus it is $.026\times \ln\frac{10^{17}}{1.5\times 10^{10}} = 0.4085$ eV below mid gap.

The Fermi level of Aluminum is 50 meV above the conduction band,
so it is $0.550+0.050 = 0.6$ eV above the midgap.

The work function difference is $-06-0.4085 = -1.0085\mathrm{V}$, $\approx -1.01\mathrm{V}$. – **[3 marks]**