ECE 314/514 Digital VLSI Design

Mid Semester Examination

3:00 PM 13-October-2020

Maximum Marks: 25 Duration: 1Hr

Instructions

1. This is a take home exam, you all should submit a single scanned pdf (high quality) in classroom, before the deadline.
2. You should definitely write your name, roll number, page number on your answer script.
3. You may choose to connect or not connect (to meting) during the quiz.
4. If you are taking any assumptions, please mention them clearly in your sheet and must make them visible distinctively.
5. Upload the whole sheet (including the rough sheet used) to avoid any plagiarism.
6. Use of calculators is allowed.
7. Institute plagiarism policy applies in this submission.
8. There is a bonus question in the end. It is over and above 25 Marks, and it's not mandatory.
9. You can ask your doubts by joining regular class link to any of us.
10. **The answer sheet must be uploaded in the classroom within 1Hr 20 Mins (before 4:20PM).**

Q1. Answer the following **[10 Marks]**

a) Describe 2 challenges of using copper as an interconnect in advanced technology nodes. [2 Marks]

b) In a long combinational path, why do we prefer to design the combinational path in early stages and use buffers in final stages. [2 Marks]

c) Calculate the standard cell height in case of a 12-track library for a technology, where metal width is 0.18um and DRC is 0.20um. Explain with the help of a diagram. [2 Marks]

d) What is the reason behind the voltage spikes coming in inverter transient simulation when the input toggles. [Hint: You can explain using a MOSFET model] [2 Marks]

e) It is observed that after fabrication, performance yield on SS lots is very low. Describe any 2 post-fabrication steps that can be taken to recover the Yield at this lot [2 Marks]

Ans: **For the questions, where we have asked to describe, some relevant reason/theory is expected. Without the description you will get only half the allotted marks.**

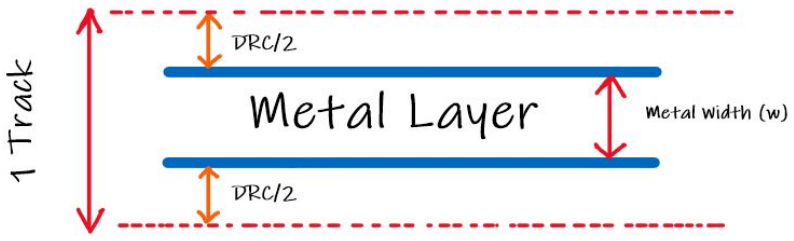
a) Challenges of using copper as an interconnect are (write any 2)
- Using copper needs barriers, and these barriers increase the resistance. Also, as scaling happens, resistance does not scale with scaling due to these barriers.
- Extreme care has to be taken so that copper does not diffuse into the silicon and disturb its band gap. Hence, extra safety is required during fabrication step.
- Dishing happens with copper, as it is a ductile material, and hence slots are required.

If other reasons are specified, that are not discussed in class, 0.5 Marks are deducted for each reason.

b) To get similar delay as an inverter, combinational gates with g>1 are sized much larger This results in increase in area, power. Also, the self-loading limits the performance gain. Therefore, such combinational gates are used early in the chain, and inverter-based buffers are used in the later part.

Those who have just written that this is done to save area, and haven't explained that why area is less when inverters are used in later stages, have been marked as incomplete and fetch 1 Marks only.

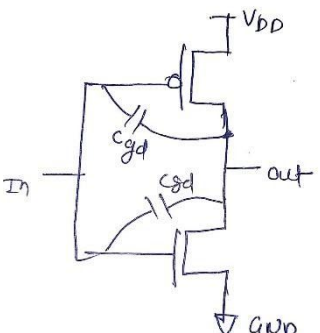
c) In case of 12 track library



1 Mark for diagram
1 Track Height = Width of metal +DRC/2 + DRC/2
= 0.18um + 0.10um + 0.10um
= 0.38um

12 Track library height is = 12*0.38 um
= 4.56 um
The above solution fetches 2 Marks.
IF someone have taken 12*metal width +11*Spacing □ this fetches only 1 Mark (with diagram) and 0 if no diagram. If the diagram is less relevant and 11*spacing answer given, then only 0.5 Marks

d) In the shown figure, we can see there is capacitance Cgd in both of the mosfets, which work as a miller coupling capacitance between the input and output.
This coupling between the input and output leads to these spikes, whenever any transition happens at the input and output. [1 Mark].
1 Mark for diagram.
Optional:
(Because of this coupling capacitance, we have two cases
1) *When input goes from 0-->1*
2) *When input goes from 1-->0.*
Explaining for the first case, when input goes from 0-->1, then, initially the voltage difference between the two plates In and



Out is VDD(out)-0(in) = VDD. When this transition is complete, due to coupling cap, the cap opposes any change in its charge and hence, tries to keep ΔV = VDD, hence, after transition when one node is at VDD, it tries Out to go to 2VDD, to make difference as VDD, and hence output goes for a rise greater than VDD. But just after this rise, the output is tied to VDD, by the devices in the inverter and comes back to VDD. This creates a spike in simulations).

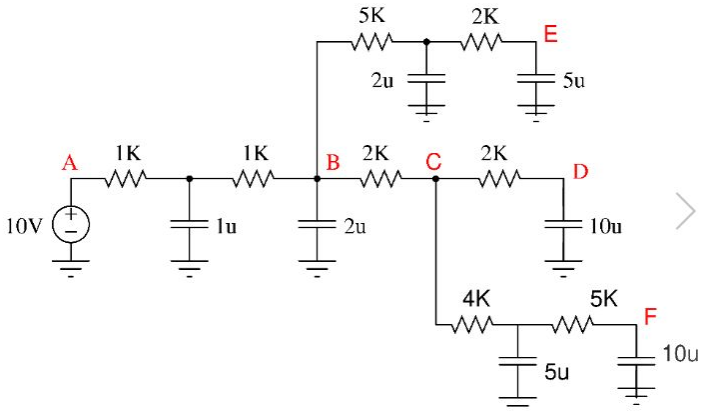
e) It is observed that after fabrication, performance yield on SS lots is very low. What can be done to recover the Yield at this lot?

The design is having yield loss at SS lot, the delay is worst in case of SS lot and this can be the reason of failing of multiple chips.
So, we would need to take steps to decrease the delay, so that now SS lot does not fail
Delay can be decreased by using these steps (Any 2 required)

1) Increase VDD, as delay decreases on increasing VDD.
2) Increase Forward biasing, to reduce Vt..(This answer fetches 1 Mark). But if only reduce Vt is given and how to reduce is not mentioned, it fetches 0.5 Marks.

If the 2 reasons are not described, just listed then 0.5 marks deducted for no explanation in each part..

Q2. Calculate the Elmore delay of path ABCF for the given RC network. **[2 Marks]**



Ans: The Elmore delay for a path takes into account

- All the resistances in the path
- And for each resistance taken, those capacitances which can be discharged or charged from current through resistor are only used

Hence, delay = 1K (1u+2u+2u+5u+10u+5u+10u) + 1K (2u+2u+5u+10u+5u+10u) + 2K (10u+5u+10u) + 4K (5u+10u) + 5K (10u)

Delay = 1K(35u) + 1K(34u) + 2K(25u) + 4K(15u) + 5K(10u)

Delay = 35m + 34m + 50m + 60m + 50m

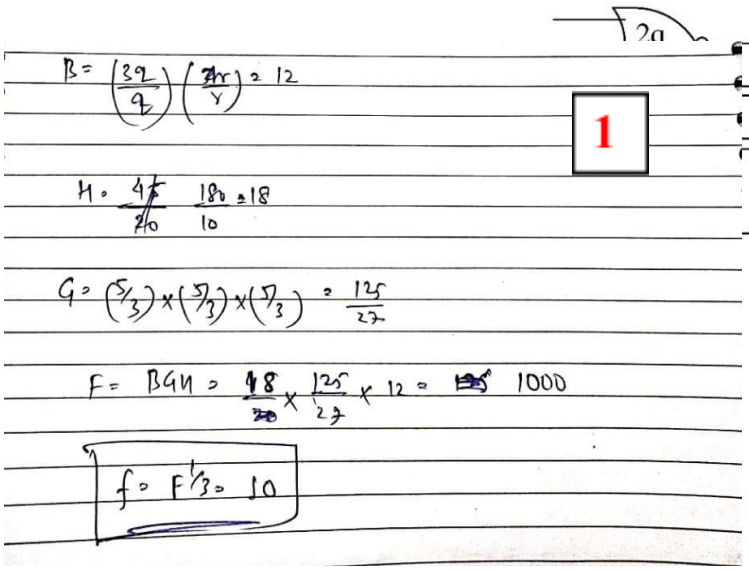
Delay = 229ms

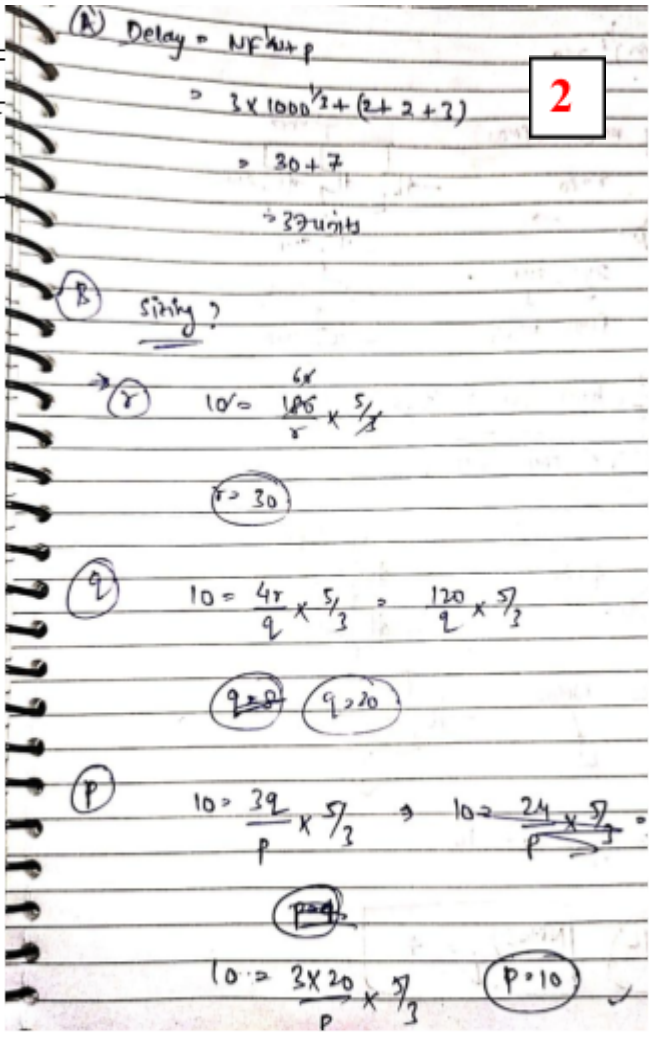
0.5 Marks deducted if unit not mentioned or wrong.

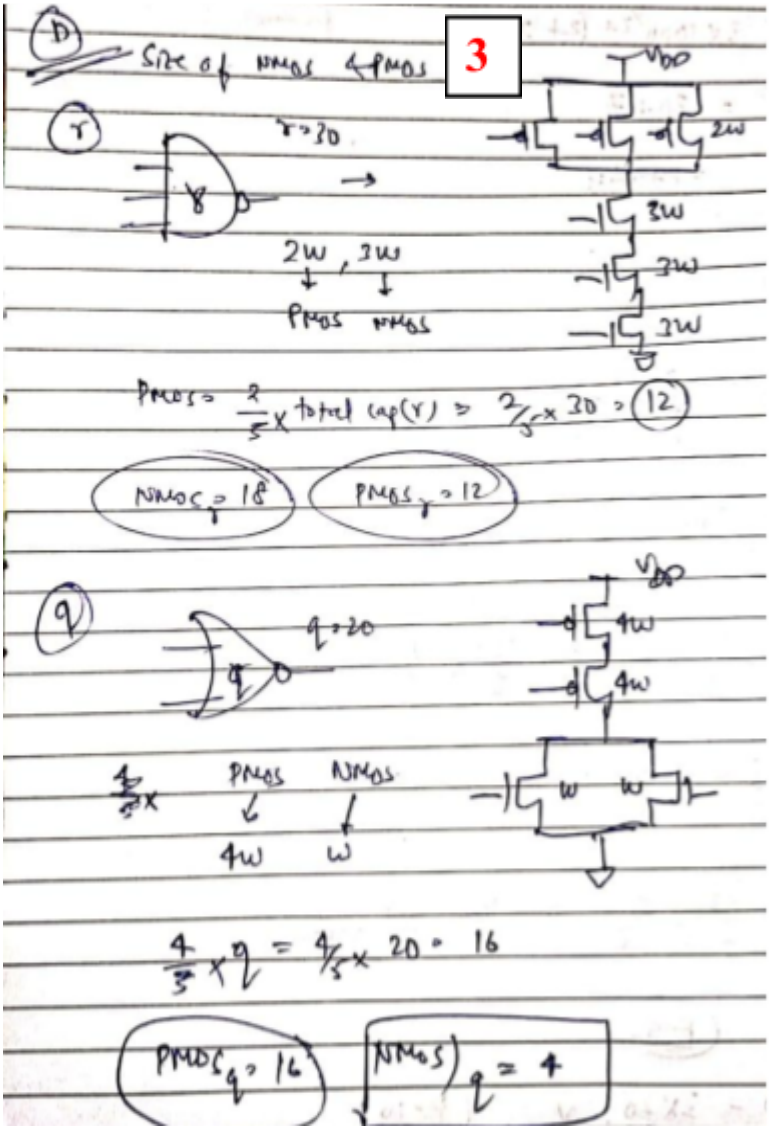
Q3. For the given design. **[8 Marks]**

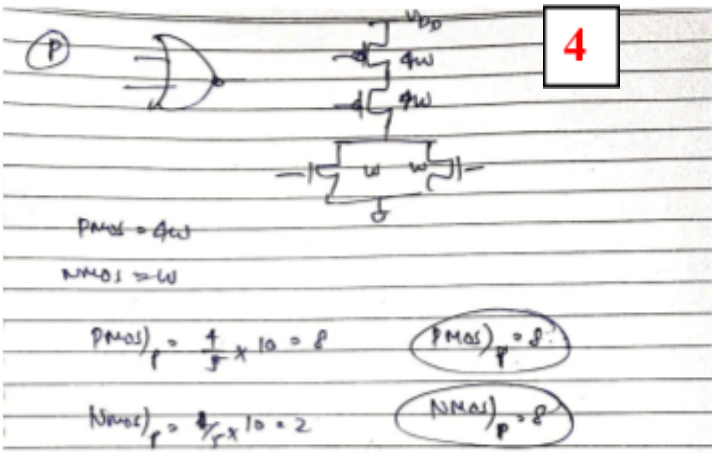
a) Minimum delay from B to E. [2 Marks]
b) Sizing for all the gates used p, q, r. [3 Marks]

c) Find the size (width) of NMOS and PMOS inside all the gates (in path). [3 Marks]







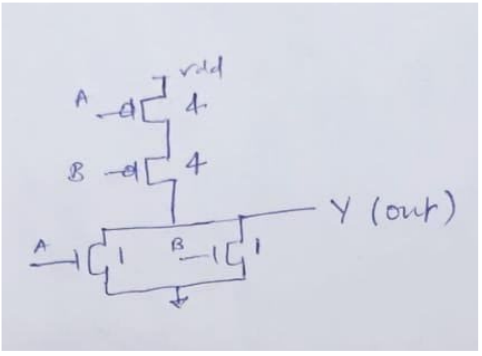


Note - If in Ques 3(a) the value of branching effort is wrong then [1 mark] for 3(a) after that if the procedure is correct in 3[b] and 3[c] then [1 mark] for 3(b) and [1 mark] for 3(c).

Q4. For a 2-input NOR gate. **[5 Marks]**

a) Size the circuit for same delay as an inverter. [1 Mark]
b) Make a stick diagram with shared diffusion. Consider, 1 finger for 2 units width of PMOS and 1 finger for 1-unit width in NMOS. [2 Marks]
c) Explain why you connected OUT, GND to specific regions in your stick diagram. [2 Marks]

Ans:



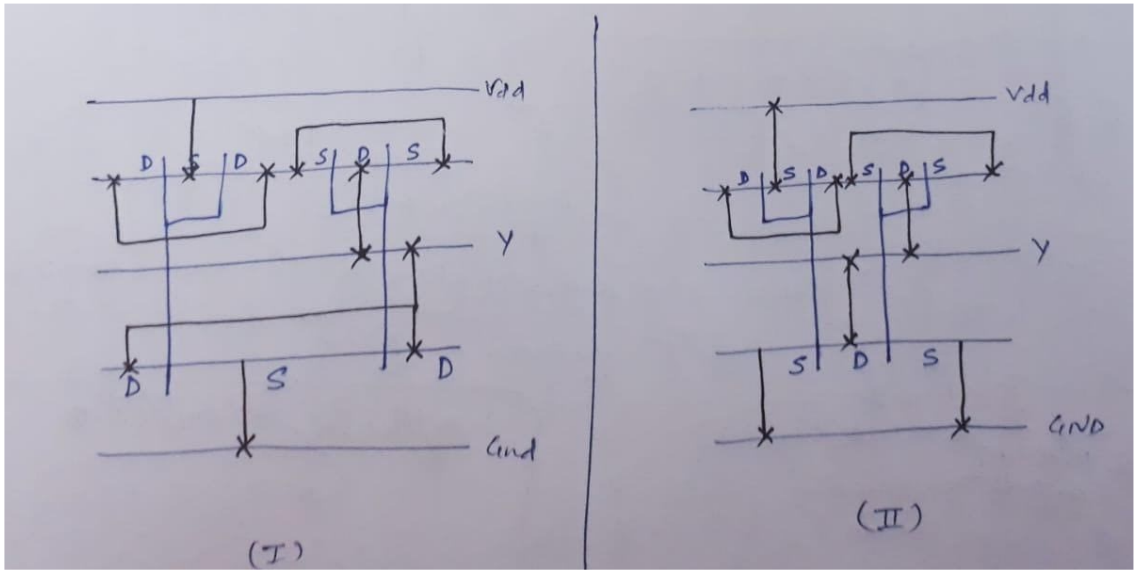
a) [1 Mark]

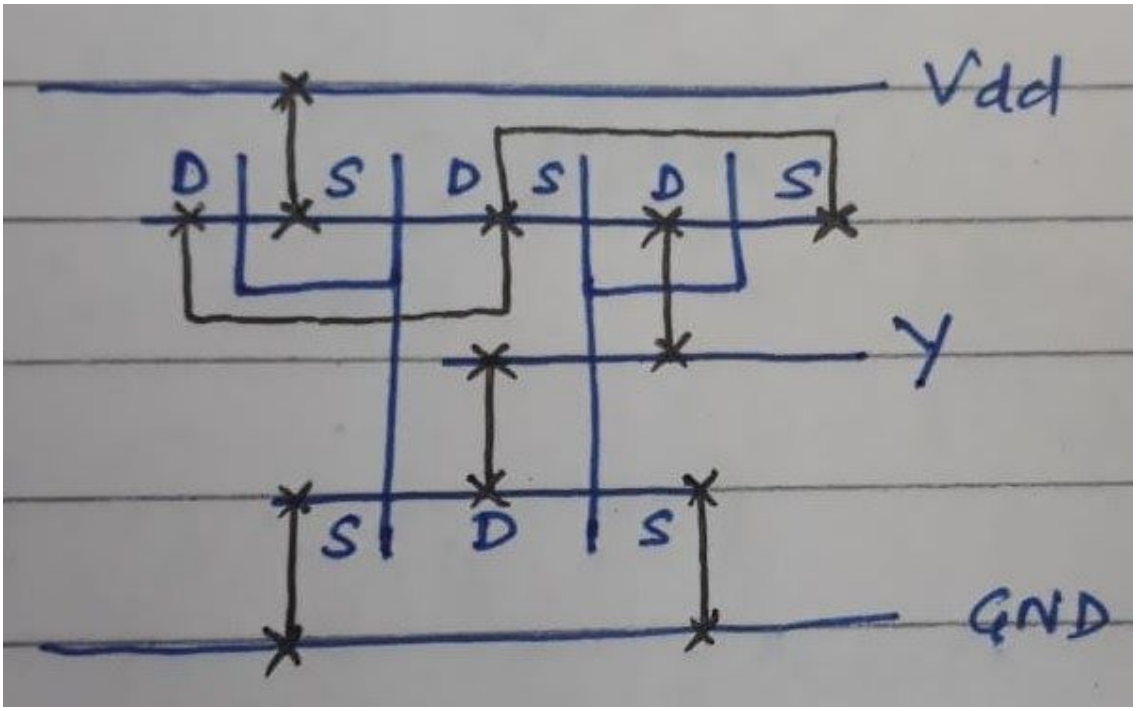
b) We are given that we have to consider 2-unit width of PMOS as one finger. Since, the width of PMOS required is 4 units, hence 2 fingers for PMOS is required and similarly, NMOS has one 1-unit width, hence can be made in one finger only.

If made stick diagram using only single finger for PMOS and NMOS, give only 0.5 Marks for correct stick diagram.

Marks in this question depends mainly on the correctness of the stick diagram. If the stick diagram is other than the 2 below figures, then also partial marks (if partially correct) may be awarded.

Given below are possible two ways of making the stick diagram.





(I) Is a correct stick diagram but not optimal [2 Marks]
(II) Is a correct stick diagram but not optimal. [2 Marks]

Reasons for (I) & (II) not being optimal are, extended diffusion region on nside, extended metals to which output is connected, multiple taps between two polys of pmos.

(III) Is the most optimal solution, where since the NMOS and PMOS are very close, hence very less diffusion width is used, which decreases the capacitance and now this less shared cap is connected to OUT. Since, OUT gets the lowest cap in this case. This is the correct answer. [2 Marks]

c) **For OUT** – We always try to put the OUT node to lowest capacitance, so as to have lesser load cap and better performance and dynamic power.

If layout made by student is (I) – Here, the diffusion is shared, but there is high diffusion capacitance as the width of diffusion here is very large. Hence, in this case, we might get lesser capacitance when using the non-shared regions for connecting output. [1 Mark]

If layout made by student is (II) - Here, the diffusion is very small, and also it is the shared diffusion, hence it is having less capacitance. Though less cap is connected to OUT but since it has multiple taps between two polys of pmos and extended diffusion and metals so its not the optimal one.. [1 Mark]

If layout made by student is (III) - Here, the diffusion is very small, and also it is the shared diffusion, hence it is having less capacitance. Hence less cap is connected to OUT. [1 Mark]

For GND – We connect GND to the higher capacitance

If layout by student is (I) then - The shared wide diffusion would have larger capacitance, as compared to the sum of two individual capacitances and hence, larger cap is connected to GND. Hence, shared is connected to GND. [1 Mark]

If Layout made by student is (II) – Here, due to less width diffusion and sharing the capacitance is lower at shared diffusion, hence the larger cap will be there in non-shared, hence connected this to GND (this also lags same points mentioned above) [1 Mark]

If Layout made by student is (III) – Here, due to less width diffusion and sharing the capacitance is lower at shared diffusion, hence the larger cap will be there in non-shared, hence connected this to GND. [1 Mark]

[Bonus Question]

Q5. Why are substrate taps required at regular frequency in a layout [1 Mark]? Explain the consequence of not using them with the help of a diagram [2 Marks]

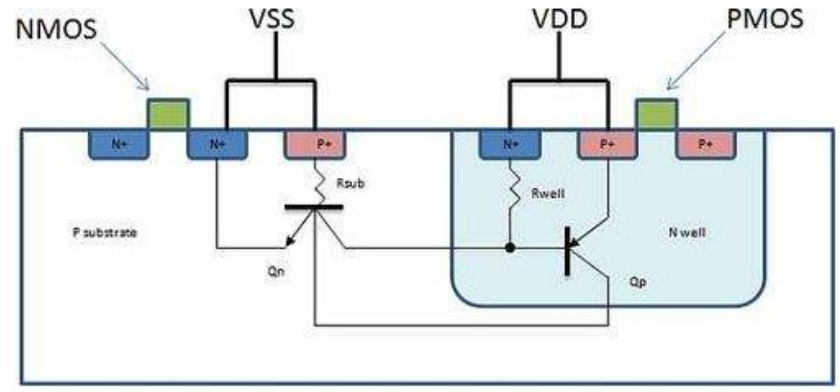
[3 Marks]

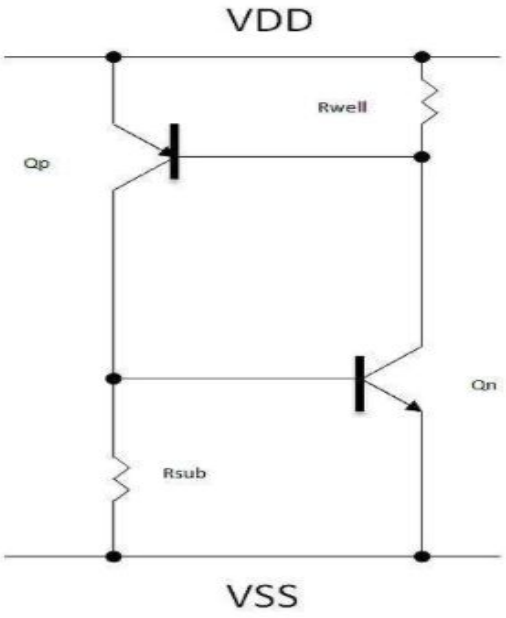
Ans:

In CMOS circuits, source-drain regions and Pwell and Nwell interact with each other to form parasitic BJT devices. These devices are connected such that they form a latch. If Rsub is very high, then huge current flows through the BJTs (like a direct VDD-GND path) resulting in chip burn-out. [1 Mark]

By putting substrate taps at regular intervals (20um specifically for 65nm CMOS), Rsub is maintained to be low to prevent Latch up. [1 Mark]

Any of these 2 figures gives 1 Mark





Using the taps at regular intervals helps to decrease the Rwell as taps provides lower resistance path to substrate to GND.

And higher substrate doping can decrease Rsub.