DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING

Spring Semester 2013-14

Answers to EEE335 Integrated Electronics 3, Questions 1…4

1. *The pull-down network for a CMOS, digital circuit is shown in* ***Figure 1****.*

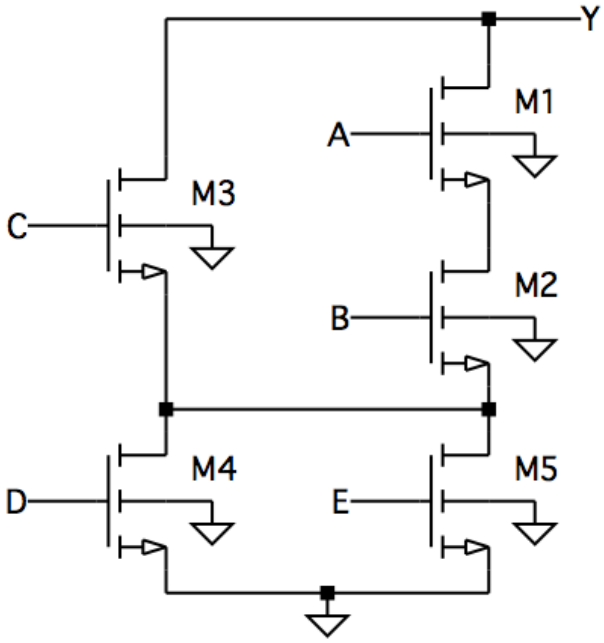
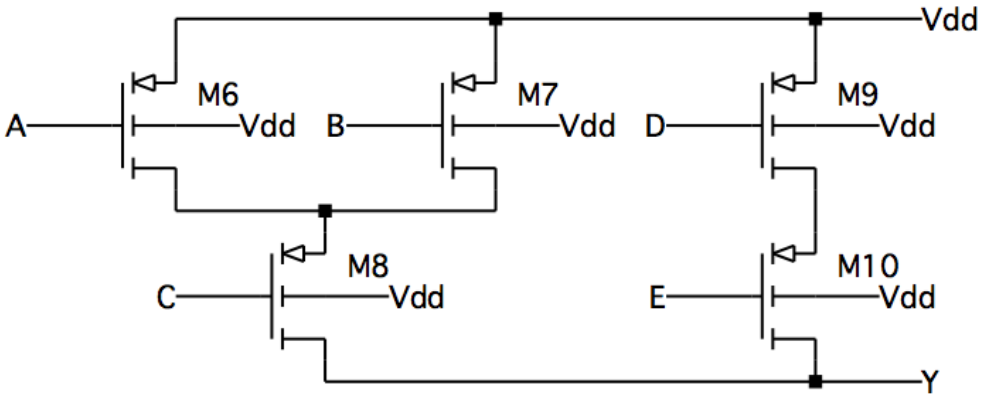


Figure 1: CMOS Pull-Down Network

a. *Draw the corresponding pull-up network for this circuit (remembering to show how the substrates are connected).*



(5)

b. *Determine the function,* ***Y****, in terms of* ***A****,* ***B****,* ***C****,* ***D****, and* ***E****.*

$Y = \overline{(AB + C)(D + E)}$ **(3)**

c. *Size the transistors* ***M1****…****M5*** *(as multiples of the minimum width of an n-type FET), assuming that the gate is 'minimum-sized'.*

Two approaches here to yield a worst case on-state resistance between the supply and the output equivalent to a minimum sized *n*-type FET – either is acceptable:

M1 = 3, M2 = 3, M3 = 3/2, M4 = 3, M5 = 3 or

M1 = 4, M2 = 4, M3 = 2, M4 = 2, M5 = 2

There is little to choose between these two solutions (one mark for each Tran.) **(5)**

d. *Why are all the substrates of these n-type FETs be connected together and to the most negative point in the circuit?*

The substrate of an *n*-type FET should be connected to a voltage equal to or more negative that the source or the drain to prevent the diodes associated with these junctions from forward-biassing. It would be helpful, therefore to connect each transistor's substrate to its source. However, this would require each transistor to be placed into a separate region of *p*-type material – isolated from the regions associated with any other transistor. This can be done but it is costly in terms of area. Alternatively, transistors can be placed in common regions of *p*-type material but must share a common substrate voltage. To ensure that this voltage is such that it does not forward-bias diodes in any transistor, it is normally connected to the most negative point in the circuit: V_{SS} or GND. **(1) (1) (1) (1)**

e. *The wire connecting the drains of **M4** and **M5** is removed. What is the new function of **Y**?*

$Y = \overline{ABE + CD}$ **(3)**

2. a. *An IC designer has to decide how to design and fabricate an ASIC:*

i) *What types of fabrication technologies are available and what are their attributes?*

Answer should include a range of approaches to implementing ASICs that are available (with some elaboration):

CPLD – few 1000 gates, deterministic behaviour/limited complexity, short design time, commodity part (10s of $), no lead time, glue logic

FPGA – up to 5M gates, non-deterministic (need full simulation), highly complex (could include large standard blocks: processors, memory, signal processing); front-end design processes follow standard 'flow', commodity parts (10-100s of $), no lead time, prototyping, low volume runs

SA – migration path from FPGA design to foundry-produced, specific IC with minimal risk for same performance. NRE in £10k+ range, lower device cost, mini, modest volume runs

MGA – pre-diffused wafers, design decomposed to NAND gates (for example) supplies metallisation, up to M gates, non-deterministic (need full simulation). Front-end and back end design process required. Foundry made but fast because wafers are standard (few weeks). NRE could be £100k+, but parts are low cost ($s) minimum order quantities e.g. 50k parts.

CBIC – from blank wafers using library of components. Full mask set required for foundry fabrication so very high NRE (up to $M). Can take up to 12 weeks to fabricate. Highest performance parts, most efficient, lowest area. Best for very big volumes but highest risk if there's a problem. **(4)**

ii) *What factors will influence the choice of technologies and the approach to design?*

The issues that affect choice will include (with some elaboration, possibly):

Size of design
Performance required
Power dissipation
Market volume

Target cost
Time-to-market
Skills/experience of design team
Any pre-existing IP (either external or internal) **(4)**

b. *Currently, the design approach often adopted by designers (at an intermediate level) is termed register-transfer level (RTL).*

i) *What does this mean and why is it important?*

RTL is a methodology that helps to ensure that designs can be implemented reliability so that they will work, as intended, when implemented in devices across the whole gamut of variation (process, temperature, voltage). Looking at a snippet of code (in VHDL):

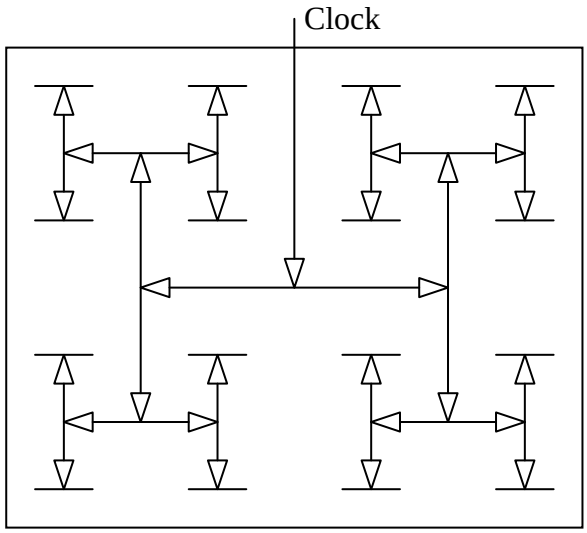
```
Process (clk)
begin
   if clk'event and (clk = '1') then
      reg3 <= reg1 OP reg2;
   end if;
end process;
```

`reg3` is a register, updated on the rising edge of the clock (`clk`) signal. Similarly, `reg1` and `reg2` are other registers written to – at the same point in time – somewhere else in the design.

With RTL, the assumption when designing, is that the value written to `reg3` on the N+1th clock cycle is the value computed from `reg1 OP reg2` using the values written to these registers at the Nth clock cycle. That is, the values written to `reg1` and `reg2` at the Nth clock cycle have a whole clock cycle to propagate through intervening logic and arrive at the register holding `reg3` in time for the correct value to by stored at this register at the N+1th clock cycle. This is predicated on the assumption that the clock signals arrive co-incidentally at all flip-flops across the IC at the same time (i.e. a signal propagating from `reg1` after the Nth clock cycle updates this flip-flop cannot arrive at the flip flop associated with `reg3` before the Nth clock cycle has updated this register). **(4)**

ii) *How might a clock tree be constructed to support the implementation of a reliable design?*

A single clock signal should be propagated across the surface of an IC by a clock network (tree) so that the delay experienced between the input of the network and any output (where flip-flops are clocked) is equal. e.g.



The clock is routed to the centre of the design, buffered, and distributed in a recursive H pattern so that each route from the input to any of the outputs is electrically the same length. The intermediate buffers allow the network to be fast by limiting the capacitance seen by any individual buffer. **(4)**

c. *Why is verification important and why is the time spent on verification beginning to dominate design activity?*

The costs of modern ASIC design and implementation can be extremely high, along with the lost opportunity cost of a device, intended for a product heading to market fails to function and results in a delay in the product entering a market. Consequently, it makes sense to verify (simulate) the behaviour and performance of the ASIC to ensure that it does what it is supposed to do, functionally, meets all the design constraints (e.g. clock frequency) and will work reliably over all process and environmental conditions. Moreover, it must also incorporate sufficient testability to ensure that malfunctioning devices can be weeded-out before they are incorporated in a product. This burden already consumes upwards of 50% of the design effort because there is a lot of verification to undertake on a large ASIC (getting correspondingly larger with larger designs) and it can be very time-consuming in its own right. However, the burden is growing, particularly as designs shrink because the an increasing number of potential problems need to be simulated to ensure that there are no problems. So, traditionally, whilst verification covered logic errors and timing, this might be extended to cover the variability of device behaviour and the variability of timing arising from signal interactions (e.g. coupled capacitance) up to partial/full electromagnetic simulations. Clearly, the time spent on verification must, at some point, be traded-off: the cost/time of verifying comprehensively .v. the risk/costs of missing errors by not simulating comprehensively. **(4)**

3. a. i)

$$V_{GS} = 10 \times \left(\frac{0.5}{4.01}\right) = 1.25V$$

(1)

$$V_{OV} = V_{GS} - V_{TO} = 1.25 - 0.6 = 0.65V$$

$$I_D = \frac{K_n}{2}\frac{W}{L}V_{OV}^2 = 40 \times 10^{-6} \times 20 \times 0.43 = 338\mu A$$

(1)

$$g_m = K_n\frac{W}{L}V_{OV} = 80 \times 10^{-6} \times 20 \times 0.65 = 1mS$$

(1)

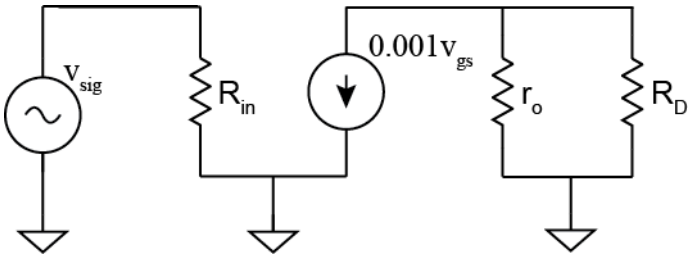
The Q-point voltage at Vout is

$$V_{out} = 5 - I_D R_D = 5 - 338\mu A \times 15k\Omega \approx 0V$$

So that V_{DS} for M1 is 5V>>V_{OV}, and the FET is in saturation.

(1)

ii)



$$R_{in} = (3.51M\Omega \parallel 500k\Omega) = 438k\Omega$$

$$r_o = \frac{1}{\lambda I_D} = \frac{1}{0.025 \times 338 \times 10^{-6}} = 118.3k\Omega$$

(3)

Ignoring channel modulation, the gain is:

$$A_v = -g_m R_D = -0.001 \times 15k\Omega = -15$$

(1)

iii)

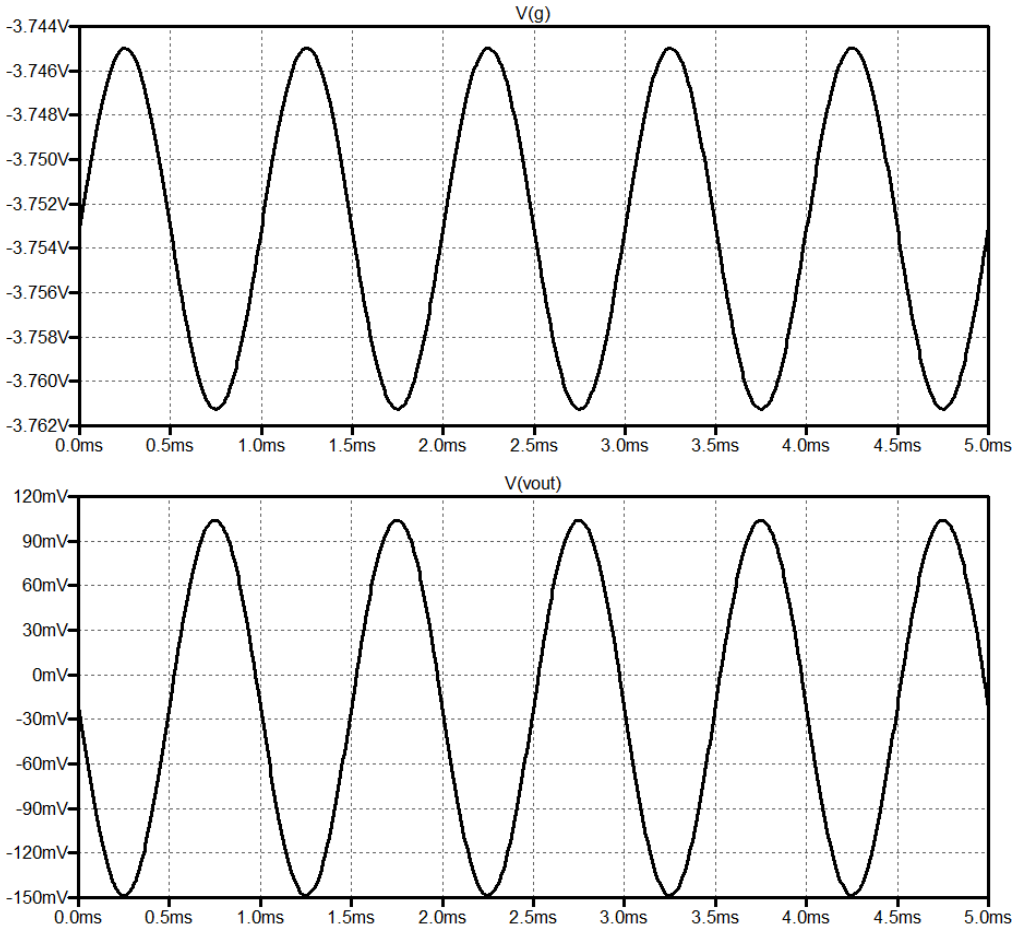
The gate-source voltage is $v_{gs} = \frac{R_{in}}{100k + R_{in}} v_{sig} = 0.814 v_{sig}$

So the gain becomes:

$$A_v = -g_m R_{out} \times 0.81 = -0.001 \times 15k\Omega \times 0.814 = -12.2$$

(2)

b.



Marks for correct offset voltages, correct sinewave magnitudes, correct x-axis labeling and inversion of output.

(4)

c.

i) The performance of the amplifier is highly dependent upon the parameters of the FET itself. These can vary between devices with the same part number, between transistors on the same chip, or as a function of temperature. The voltage divider biasing holds the gate voltage constant regardless of the actual current that flows through the FET, and so the amplifier's Q-point and gain are both unpredictable. Better biasing requires some form of feedback, so that variations in the drain current of the FET are controlled by appropriate variations in V_{GS}. **(3)**

Voltage divider biasing is unsuitable for implementation on an IC because large value resistors take up huge amounts of room on a semiconductor substrate. The coupling capacitor also takes up valuable chip area. A better solution is to use a self-biased differential amplifier that doesn't require any capacitive signal coupling.

ii.) Since the curve of I_D vs V_{DS} for M2 in the diagram above is proportional to the reciprocal of the output resistance of the mirror, maximizing this resistance means that V_{DS}(M2) can vary over a large range of voltages (as may be the case when the mirror is attached to a circuit), without a significant change in the current it pulls. When used for amplifier biasing, this means that Q-points within the amplifier can be made extremely stable, since the current mirror will automatically set a value of Vout that delivers the designed bias current over a large range of operating conditions. **(3)**

4. a. i)

$$V_{OV} = \frac{1}{\lambda\sqrt{2500}} = \frac{1}{0.05 \times 50} = 0.4V$$ **(1)**

For FETS M3,4,5,6:

$$r_{o1} = \frac{2}{\lambda I_{ref1}} = \frac{2}{0.05 \times 150 \times 10^{-6}} = 266.7k\Omega$$ **(1)**

For FETs M7,8:

$$r_{o2} = \frac{1}{\lambda I_{ref2}} = \frac{1}{0.05 \times 300 \times 10^{-6}} = 66.7k\Omega$$ **(1)**

ii)

The transconductance of the differential amplifier is:

$$g_{m1} = \frac{150 \times 10^{-6}}{0.4} = 0.375mS$$

Whilst that of the CS amplifier stage is:

$$g_{m2} = \frac{600 \times 10^{-6}}{0.4} = 1.5mS$$ **(1)**

The gain of the differential pair is then:

$$A_{diff} = g_{m1}(r_{o1} \parallel r_{o1}) = \frac{0.000375 \times 266700}{2} = 50$$ **(1)**

The gain of the CS amp is:

$$A_{CS} = g_{m2}(r_{o2} \parallel r_{o2}) = \frac{0.0015 \times 66700}{2} = 50$$ **(1)**

b. To find the channel aspect ratios:

For M1:

$$300\mu A = \frac{1}{2} K_p \frac{W_1}{L_1} V_{OV}^2$$

$$300\mu A = \frac{1}{2} 150 \times 10^{-6} \times \frac{W_1}{L_1} 0.4^2$$

$$\frac{W_1}{L_1} = 25, \quad W_1 = 18.75\mu m$$ **(1)**

M2 must be chosen to provide I_{ref1}=150µA, i.e half the master current, I1, of the current mirror, so:

$$\frac{W_2}{L_2} = 12.5, \quad W_2 = 9.38\mu m$$ **(1)**

For M3,4:

$$\frac{I_{ref1}}{2} = \frac{1}{2} K_p \frac{W_{3,4}}{L_{3,4}} 0.4^2$$ **(2)**

$$\Rightarrow \frac{W_{3,4}}{L_{3,4}} = 6.25, \quad W_{3,4} = 4.69\mu m$$

For M5,6:

$$\frac{I_{ref1}}{2} = \frac{1}{2} K_n \frac{W_{5,6}}{L_{5,6}} 0.4^2$$

$$75\mu A = \frac{1}{2} 200 \times 10^{-6} \times \frac{W_{5,6}}{L_{5,6}} 0.4^2$$ **(1)**

$$\Rightarrow \frac{W_{5,6}}{L_{5,6}} = 4.69, \quad W_{5,6} = 3.52\mu m$$

For M7:

$$I_{ref2} = \frac{1}{2} K_n \frac{W_7}{L_7} 0.4^2$$

$$\Rightarrow \frac{W_7}{L_7} = 18.75, \; W_7 = 14.06\mu m$$ **(1)**

For M8:

$$I_{ref2} = I1 \frac{W_8}{L_8} \Big/ \frac{W_1}{L_1}$$

$$\Rightarrow \frac{W_8}{L_8} = \frac{I_{ref2}}{I1} \frac{W_1}{L_1} = \frac{300}{300} 25 = 25$$

$$\Rightarrow W_8 = W_1 = 18.75\mu m$$ **(1)**

c. A virtual ground exists at the drain of M2, or at the combined source of M3 & M4. **(1)**

Since the op-amp is driven by a balanced load with no common-mode component, the Q-point voltage at this point is $0V+V_{TO}+V_{OV}$=0.9V. **(2)**

d. The Miller multiplier is equal to 1 + the gain of the CS amp = 51. The capacitor **(4)**

C1 appears in parallel with C_{GD}=10fF, so that the total Miller-multiplied capacitance is:

$$C_M = 51 \times (0.02 \times 10^{-12} + C1)$$

The total capacitance at the gate of M7 is then:

$$C_{tot} = C_M + C_{GS} = 51 \times (0.02 \times 10^{-12} + C1) + 0.05 \times 10^{-12} \qquad (*)$$

The resistance seen by C_{tot} is $r_{o1} \parallel r_{o1}$=133.3kΩ, so the upper cutoff frequency is approximately:

$$f_{cutoff} \approx \frac{1}{2\pi \times 133333 \times C_{tot}} = 100kHz,$$

$$C_{tot} \approx \frac{1}{2\pi \times 133333 \times 100000} = 11.9pF$$

Rearranging * to find C1:

$$C1 = \frac{C_{tot} - 0.05 \times 10^{-12}}{51} - 0.02 \times 10^{-12} = 0.21pF$$

NLS AMM