Written examination in **Integrated Circuit Design MCC091/MCC090**

Wednesday August 26, 2013, at 14.00-18.00 in the Maskin building. ☺

Staff on duty: Kjell Jeppson, MC2, phone ext: 1856, or mobile 0703-088581, and Lena Peterson 1822.
Administration: Send exams to Kjell Jeppson, MC2/BNSL, and send lists to Jeanette Träff, MC2/Photonics.

Technical aids for students: This is a closed-book examination, only the design rule sheet and a calculator is allowed.
The results from the examination will be sent to you via the Ladok system within three weeks. The exact time and place for review of this examination will be posted later on the course web page. Solutions will be posted at the course web site in pingpong after the examination is over.

==========

The written examination contains six problems, each worth 10 points. You need 30 points to pass, 40 points for grade "4" and 50 points for grade "5".

==========

1. **Some warm-up questions:**

a) **What** is the difference between a contact and a via? (2p)

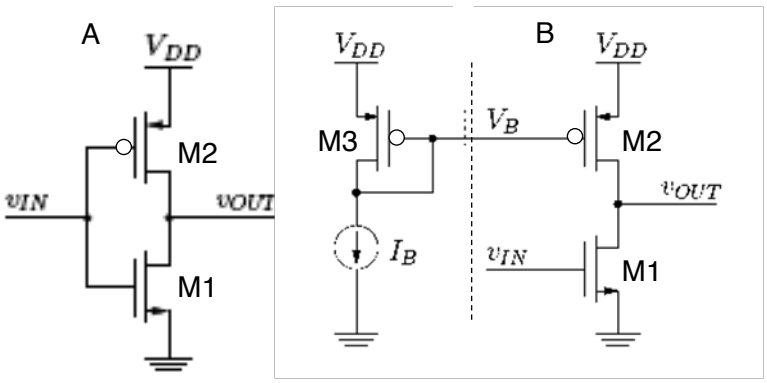
b) **Explain** why the SiO_2 has been replaced by a high-κ insulator in most recent technology nodes! (2p)

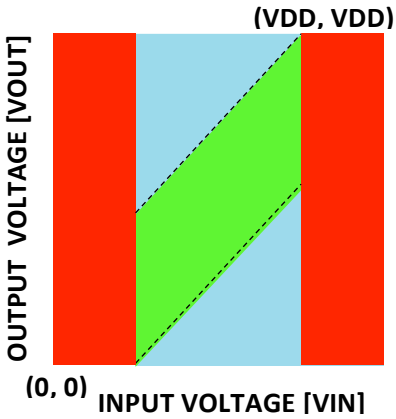
c) **What** is the reason for using $0.7C_LV_{DD}/I_{DSAT}$ to calculate the logic gate propagation delay instead of using the step response delay expression $0.5C_LV_{DD}/I_{DSAT}$? (2p)

d) **What** does the logical effort of a logic gate say about a gate with respect to the inverter? (2p)

e) **How** are branches handled when calculating the Elmore delay for a clock tree? (2p)

2. **MOSFETs and inverter voltage transfer curves**





In the figure above you find the circuit diagram for two CMOS inverters, (A) and (B). The inverter to the left (A) is an ordinary CMOS inverter that flips at $V_{DD}/2$. The inverter to the right (B) is a pseudo-NMOS inverter intended as an amplifier. Its load p-channel MOSFET M2 is biased at an unknown gate voltage V_B determined by a current mirror. A current mirror takes current I_B from a constant-current source and mirrors it to the inverter. Otherwise, the two CMOS inverters are identical except for the biasing. Transistors M1 and M2, respectively, are the same MOSFETs in both inverters. The rightmost diagram in the figure above shows the MOSFET regions operation in CMOS.

a) **Relate** the two current gain factors k_1 and k_2 of MOSFETs M1 and M2 to each other considering that inverter (A) flips at $V_{DD}/2$ assuming symmetrical threshold voltages, $V_{TN}= -V_{TP}=V_{DD}/5$? (2 p)

b) Assuming $V_B=0.6V_{DD}$, **what is** the flipping voltage of the pseudo-NMOS-inverter (B)? (2 p)

c) **Calculate** current I_B if $V_{DD}=1.2$ V, and $k=600$ µA/V^2! (2 p)

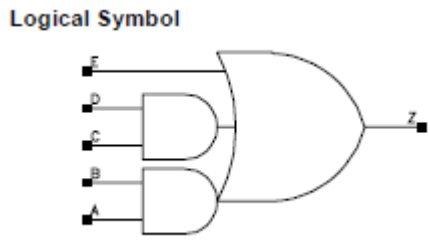
d) For **what** output voltage range are both MOSFET devices saturated in inverter (A)? (2 p)
Refer to the right-hand diagram showing the CMOS regions of MOSFET operation!

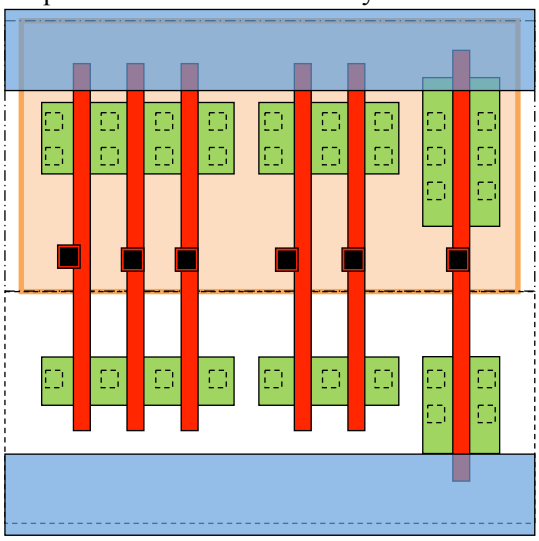
e) For **what** output voltage range are both devices saturated in the pseudo-NMOS inverter? (2 p)
Refer to the right-hand diagram showing the CMOS regions of MOSFET operation!

3. Multi-core. One of the reasons for going to multi-core computer systems is to reduce the total power by lowering the supply voltage and then compensate for speed losses by using more than one core that work in parallel. A simplified assumption is that four cores at a fourth of full speed can do the same work as one core at full speed.

From these simple assumptions and from the assumption that the delay is given by CV_{DD}/I_{DSAT}, find out **how much** we can reduce the power dissipation by using a quad-core design and while still getting the same job done? Assume that we use the same CMOS process and the same core, but that 20% of the core capacitance must be added to the total chip capacitance for the control unit coordinating the four cores. Let us assume that V_T is 25% of the original supply voltage. Use a simple square-law current model for the MOSFET saturation current! (10 p)

4. CMOS layout: Use the enclosed layout template to **draw** the metal1 layout of an AO212 gate! (10 p)

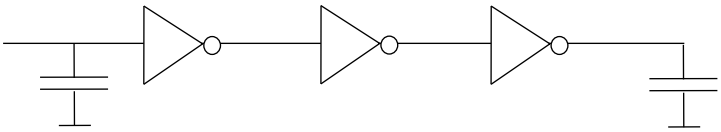




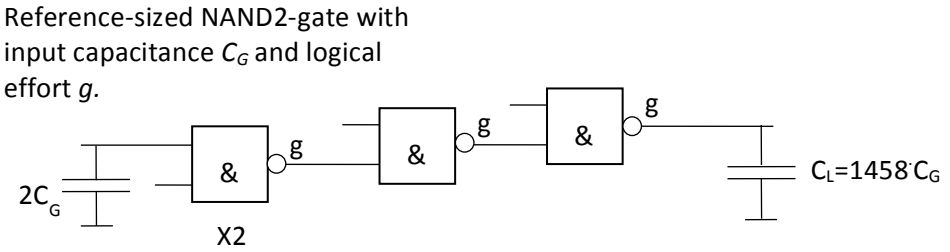
2+2+1 AND-OR (AO212) gate

5. Tapering, logical effort and branching.

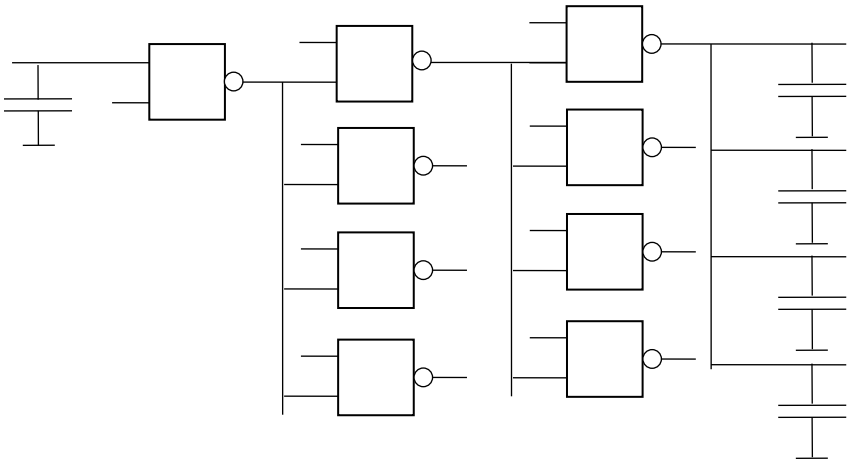
a) Assume that we are building a buffer driving a large capacitive load. **What sizes** should be chosen for the buffer inverters to minimize the delay if the first inverter is of size X2? (4 p)



b) Assume that we replace the inverters with 2-input NAND-gates to implement also a clock gater function. **What sizes** should be chosen for the buffer NAND gates to minimize the delay if the first NAND gate is of size X2? (4 p)



c) **How** would branching change your design considerations? (2 p)

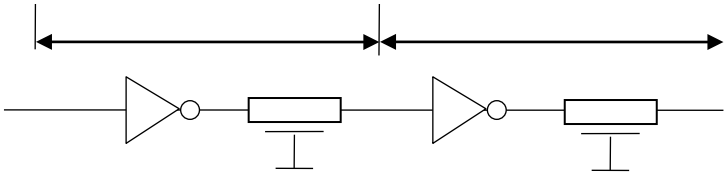


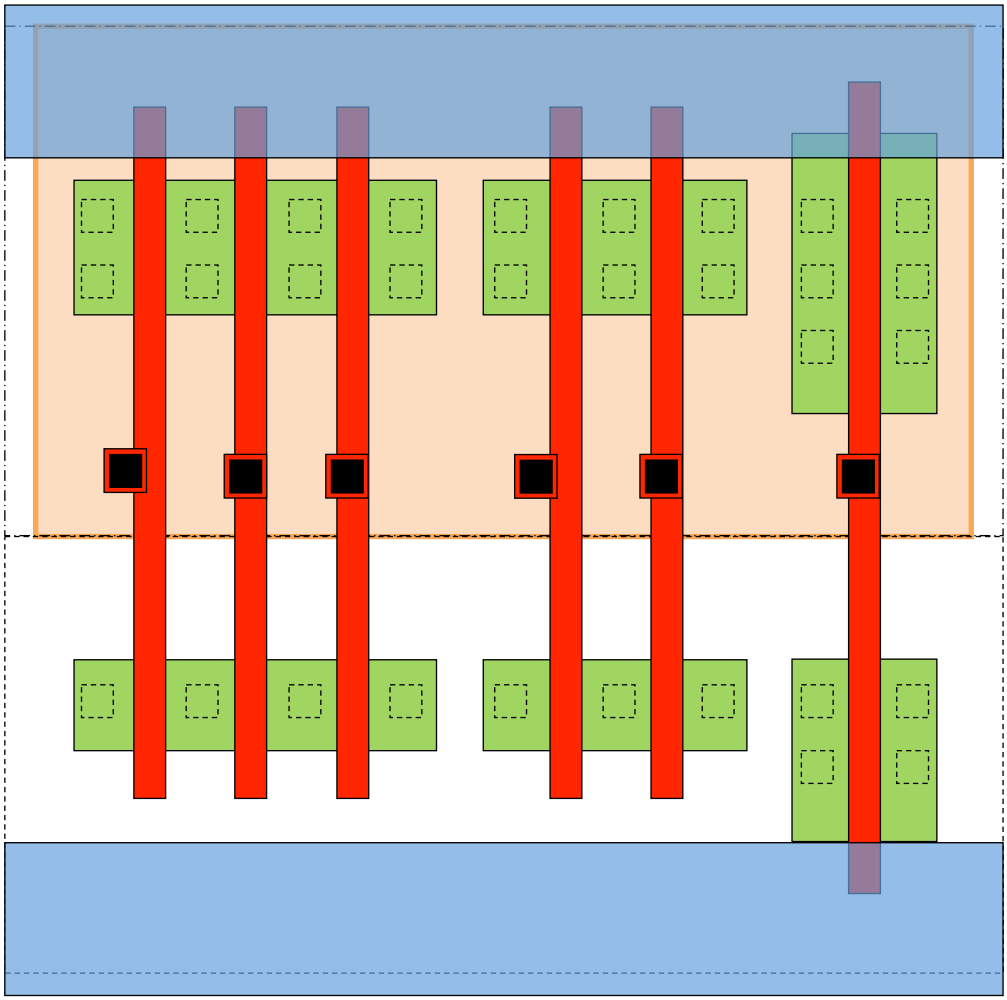
6. Wire delay and repeaters

A bus wire from the data path of a processor to the on chip data cache memory has the resistance R_w, and the capacitance C_w. The intrinsic inverter delay for this technology node, i.e. the product of the inverter resistance and the inverter input capacitance, is equal to RC.

a) Use the Elmore delay model to **derive an expression** for the number of repeaters that should be inserted to minimize the bus delay! Neglect the parasitic inverter output capacitance. (5 p)

b) **What would be the best value** for the repeater resistance R with respect to the wire segment resistance, R_w/m? (5 p)





This sheet can be detached and handed in with your layout solution. Exam code number:_____

1. Short answers to warm-up questions

a) A contact is made from the first metal layer to any of the two front-end layers below, contacts are between metal one and either active or poly. A via is a connection between back-end metal layers, i.e. between metal layer *n* and metal layer *n*+1.

b) A high-κ insulator layer can be thicker and still provide a high C_{ox}.
With a thicker gate insulator layers the gate leakage current is reduced.

c) 40% extra delay is added to account for a typical input voltage not being a step function (but a signal with a typical rise/fall time equal to the output fall/rise time).

d) The logical effort relates the gate RC product to the inverter RC product.

e) Branch capacitances are short-circuited to the branching point along the main critical path.

2. Amplifiers and inverters

a) They must be equal, $k_1=k_2$, since $V_{sw}=V_{DD}/2$.

b) It flips when both devices have the same gate voltage overdrive, i.e. $V_{IN}=V_{DD}-V_B=0.4V_{DD}$.

c) The current through the PMOS load is given by

$$I_{DS}=\frac{k_2}{2}\left(V_{DD}-V_B+V_{TP}\right)^2=300\left(1-0.6-0.2\right)^2 1.2^2=17\mu\text{A}.$$

d) The region where both M1 and M2 are saturated is given by the saturation conditions, i.e.: $V_{sw}-V_{TN}<V_{OUT}<V_{sw}-V_{TP}$, i.e. $0.3V_{DD}<V_{OUT}<0.7V_{DD}$.

e) $V_{sw}-V_{TN}<V_{OUT}<V_B-V_{TP}$, i.e. $0.2V_{DD}<V_{OUT}<0.8V_{DD}$.

3. **Multi-core**. *In this problem students can demonstrate basic knowledge about simple delay and power dissipation formulas and concepts.* The propagation delay in a processor is given by

$$t_d:\ \frac{CV_{DD}}{I_{DSAT}},\ \text{where } I_{DSAT}=\frac{k}{2}\left(V_{DD}-V_T\right)^2.$$

With four processors we assume that the frequency of operation can be reduced to one quarter of the original frequency. Hence, the delay can be increased by a factor of four by reducing the supply voltage to the *x* fraction of the original supply voltage V_{DD}. Equaling the two delay expressions yields the following equation:

$$4\cdot\frac{1}{\left(1-0,25\right)^2}=\frac{x}{\left(x-0,25\right)^2},$$

Solving this equation yields x=0.5 V as shown below:

$$4\cdot\left(x-0,25\right)^2=x\left(\frac{3}{4}\right)^2\Rightarrow 4\left(4x-1\right)^2=9x\Rightarrow 64x^2-32x+4=9x$$

$$\Rightarrow 16x^2-10x+1\approx 0\Rightarrow x=\frac{5}{16}\pm\frac{\sqrt{25-16}}{16}=0.5$$

The dynamic power dissipation is given by

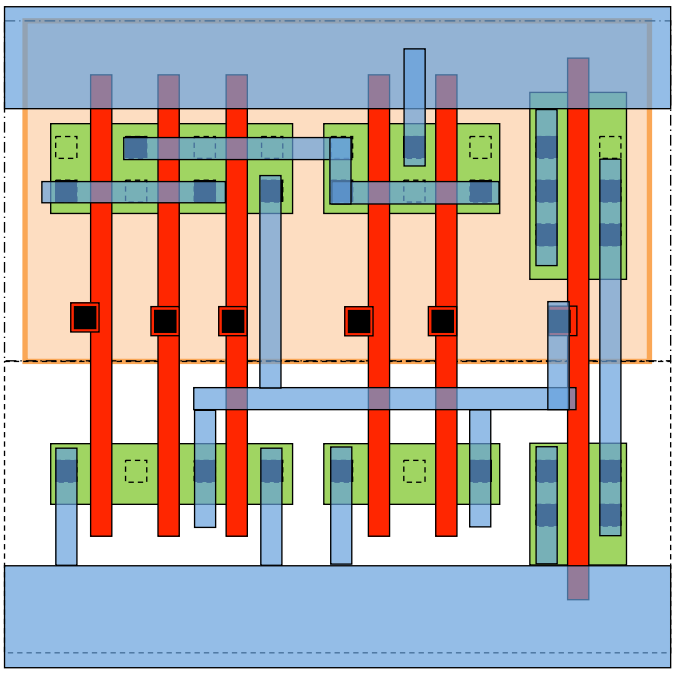
$$P_d=fCV_{DD}^2.$$

Insertion of new numbers for frequency, capacitive load, and supply voltage yields

$$P_d\approx\left(\frac{f}{4}\right)\cdot\left(4.2C\right)\cdot\left(\frac{V_{DD}}{2}\right)^2\approx 0.25fCV_{DD}^2.$$

Hence, the quad-core dynamic power dissipation can be reduced to one quarter of the single core power dissipation. At least in an ideal world! ☺

4.



5. **Tapering, logical effort and branching.**

a) The path electrical effort is given by H=1458/2=729=9^3. Hence, f=$\sqrt[3]{729}$=9. The inverter sizes are X2, X18, and X162.

b) The logical effort of a 2-input NAND is g=4/3. The path logical effort is G=ggg=64/27.

The path fanout becomes F=GH=1728=12^3. Hence f=12. For each gate h=f/g=9. The NAND sizes become X2, X18, and X162 as before.

c) The path branching is B=$b_1b_2b_3$=4^3. F=GHB=12^3*4^3. Hence, f=gh=48. For each gate h=36. Since there are four identical NAND gates loading, we obtain a scaling factor of 9 as before. The NAND gate sizes remains X2, X18, and X162 as before.
Checking: 9*4*4/3=48. OK! But it is quite heavy loads for these NAND gates, maybe there should be 4 NAND gates in the buffer instead of only 2.

6. **Wire delay and repeaters**

a) Divide the wire into m sections, each with a resistance R_w/m and a capacitance C_w/m. Insert m repeaters, one for each wire section. Using the Elmore delay model we obtain the following expression for the delay:

$$delay = m\left(R\left(C+\frac{C_w}{m}\right)+\frac{R_w}{m}\left(C+\frac{C_w}{2m}\right)\right).,$$

Taking the derivative wrt m, we obtain:

$$RC=\frac{R_wC_w}{2m^2}\Rightarrow m=\sqrt{\frac{R_wC_w}{2RC}}.,$$

b) Taking the derivative wrt to R (remembering that RC is a constant), we obtain

$$C_w=\frac{R_wRC}{R^2}\Rightarrow R^2=R_w{}^2\frac{RC}{R_wC_w}=\left(\frac{R_w}{m}\right)^2\frac{m^2RC}{R_wC_w}=\left(\frac{R_w}{m}\right)^2\Big/\sqrt{2}.$$