Name _____

Name: _____
I acknowledge this exam has been taken under all aspects of the ND Honor Code.

CSE 462 VLSI Design: Final Exam
Dec. 12, 2018

- SIGN YOUR NAME!!!!!!
- Open book and notes, but no computer searches, cell phones, or communications with or help from others. The use of a computer is permitted ONLY for calculations or access to the course web site. All aspects of the ND Honor code apply.
- Do all problems. Show all you work in the spaces supplied, including equations that you may have used.
- Show units on your final answers to numerical questions.

For reference:

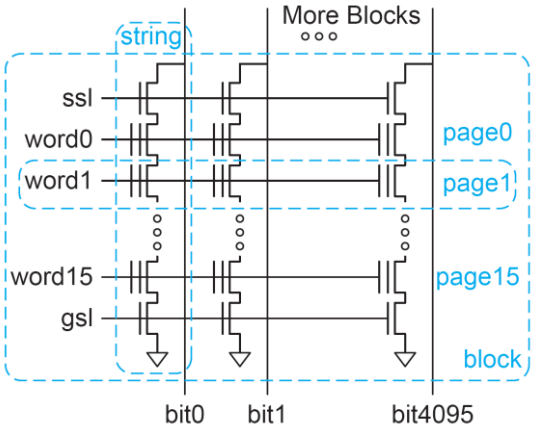
- 1us = 10^{-6} seconds
- 1ns = 10^{-9} seconds
- 1ps = 10^{-12} seconds
- 1um = 10^{-6} meters
- 1nm = 10^{-9} meters
- 1mm^2 = 10^6um^2 ~ 2^{20} um^2
- 1 GHz = 10^9 Hz
- 1 joule = 1 watt x 1 second
- 1 joule = 1 Farad x 1 volt^2
- 1 KΩ = 10^3 ohms
- 1 nF = 10^{-9} Farads
- 1 pJ = 10^{-12} joules
- 1 fF = 10^{-15} Farads = 10^{-6} nF
- 1 KB = 2^{10} = 1024 bytes (inconsistent with K as in ohms)
- 1 MB = 2^{20} = 1.048x10^6 bytes
- 1 GB = 2^{30} = 1.073x10^9 bytes
- Note: 1 nF x 1 GHz x1V^2 = 1 Watt

1	Multiple Choice	20	
2	Scaling	20	
3	Multi-gate Delay	20	
4	Gate Delay	20	
5	IV	20	
	Total		

Gate Type	Number of inputs														
	1			2			3			4			N		
Inverter	3	1	1												
NAND				4	2	4/3	5	3	5/3	6	4	6/3	N+2	N	(N+2)/3
NOR				5	2	5/3	7	3	7/3	9	4	9/3	2N+1	N	(2N+1)/3
TriState/mux		2	2		4	2		6	2		8	2		2N	2
XOR, XNOR					4	4,4		6	6,12,6		8	8,16,16,8			

(x,y,z) = Input Cap, p, g

Name _____



1. (20 pt) Multiple Choice/Short Answer. You do not need to multiply out numbers, and feel free to list answers in powers of 2

a) Consider the flash memory block from Fig. 12.60 p. 531 (reproduced aside). Assume each cell can store 8 levels. How many bits of information can be saved in this block? Show math

b) _____The logical effort for a 3 input NOR is what?

c) _____What is the normalized delay for a 2 input NAND of minimum size driving a load of 36C

d) _____For a NAND2 what is the width of the n-types that would give a drive of 6.

e) _____Using Example 4.11 as a reference, what is the frequency of a ring of 9 65nm inverters that have double their normal widths?

f) _____Assume that the delay through some circuit is 36RC, where R and C are nominal on resistance and input capacitance of a unit inverter. What is the normalized delay?

g) _____What is the load capacitance on a NOR3 with unit size n-type transistors and a normalized delay of 17.

h) _____The FeFET transistors that Dr. Niemier discussed have what properties (list all that match). (a) are just like CMOS, (b) remember their state when they are powered down, (c) use magnetic fields for storage, (d) cannot be used for memory, (e) may make for dense neural nets, (f) none of above

i) _____The floating gate of a flash memory transistor (A) is used to activate the transistor, (B) is above the regular gate, (C) can only hold 2 levels of charges, (D) is connected to the drain, (E) none of the above.

j) _____What is the dynamic power of a gate where the activity factor is 0.5, the capacitance is 6pF, V is 2 volts, and the clock is 3GHz?

Name _____

2. (20 pt) The original Google TPU chip had a 256x256 array of 8bit multiply accumulates (MACs) (integer only, your project with 8b floats is more like TPU2). The array can perform a matrix by matrix multiply where each matrix is 256x256 in a "pipelined" fashion. The characteristics of this array on the TPU chip is given in the first column of the table below (numbers have been adjusted a bit for simplicity).
   a) Fill in the following table assuming the shrink from 28 to 14nm is Dennard scaling and the shrink from 14 to 7 is Constant Voltage. To simplify life, where possible, leave as numbers in terms of powers of 2.

Feature Size	**28 nm**		**14 nm**		**7nm**	
Vdd	**1 V**		**0.5V**		**0.5V**	
	256x256 array	One MAC	256x256 array	One MAC	256x256 array	One MAC
Area	80 mm^2	$10x2^{-13}$ mm^2				
Power	16W					
Clock	800MHz					
Power Density	$=16/80 = 0.2W/mm^2$					
Peak performance for 256x256MACs	$2^{16}x0.8x10^9 = 0.1x2^{19}$Gmacs/sec Or approx. 50 Trillion macs/sec					
# MACs that fit in 80^2mm	$256x256 = 2^{16}$					

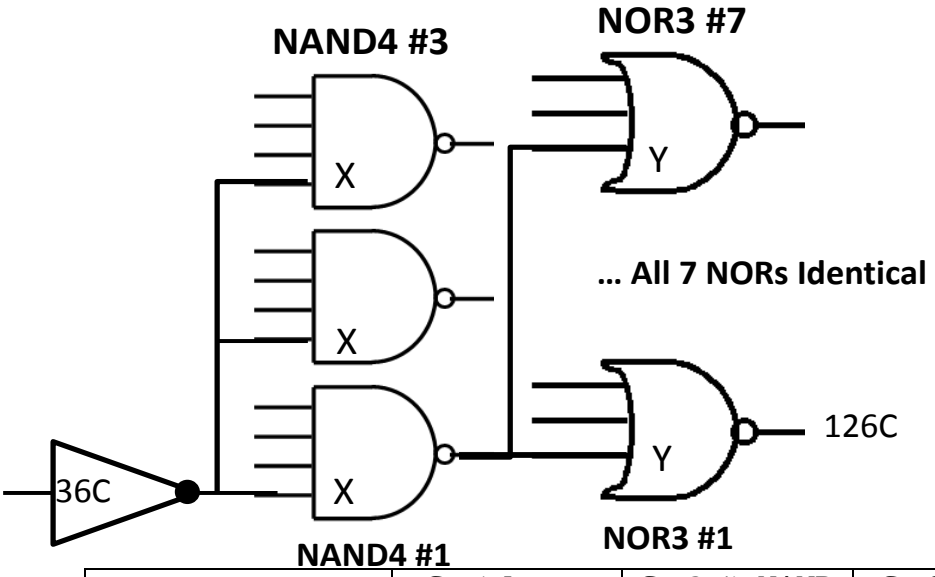
   b) _____ If each MAC was square, how long in mm, um, and nm is one side??

   c) _____ If each MAC was square, what would be its length in lambda?

   d) _____Assuming the wiring track method of sizing was accurate, how many wiring tracks in each direction?

Name _____

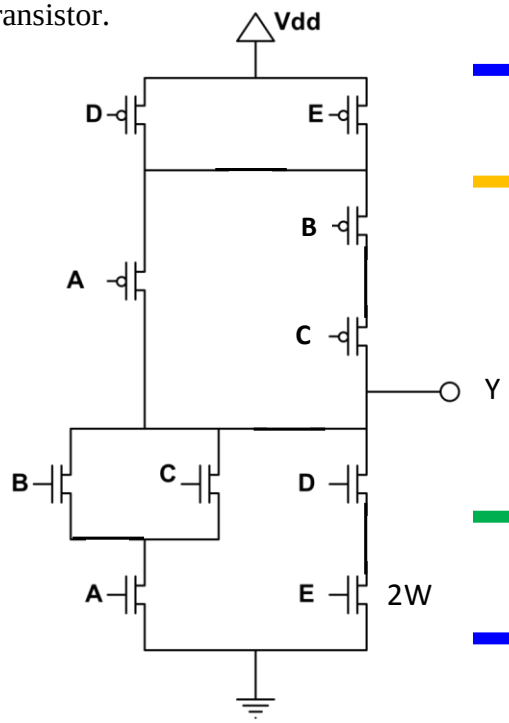
3. (20pt) Consider the following circuit consisting of an inverter driving **3 NAND4s**, with the output of the one on the path we are interested in driving **7 NOR3s**. Assume a load of 126C and an aggregate input capacitance on each input of each stage as 36, X, and Y as shown. Using the logical effort approach, compute the terms below to find transistor sizes that give an optimal delay. There are unused rows at the bottom of the table you are "encouraged" to use to track intermediate results. You are doing things right if the final numbers are whole integers. Show work!

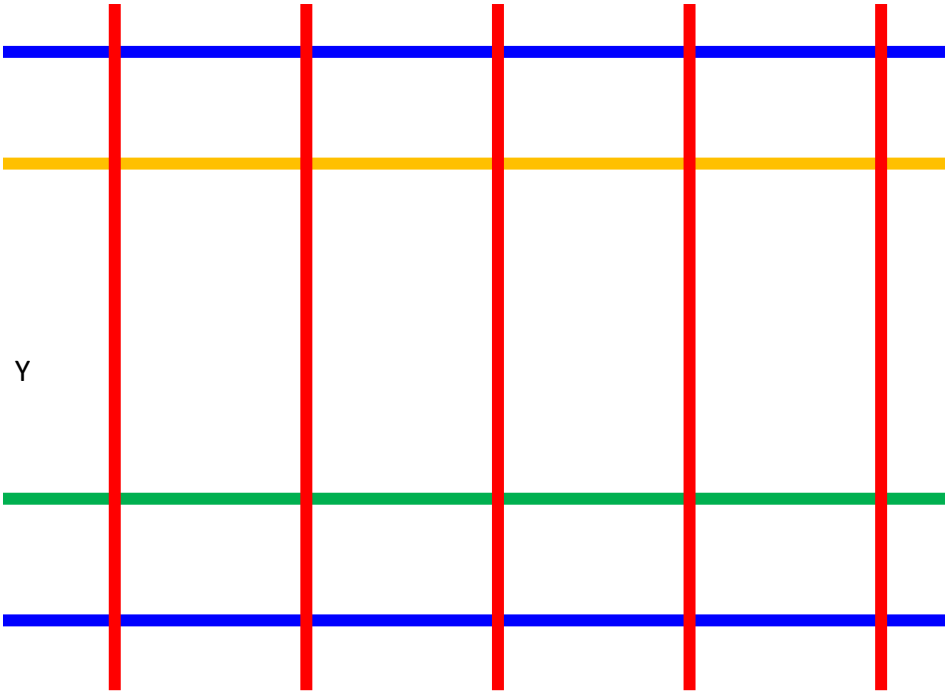


	Gate1: Inverter	Gate2: 4in NAND	Gate3: 3in NOR	
p (use table 4.3)				**P=**
g (use table 4.2)				**G=**
b			**1**	**B=**
GBH		**f^**		**H=**
Input Capacitance	**36**	**X=**	**Y=**	
Scale Factor on Transistors				
Final Width of P Type				
Final Width of N Type				
Normalized Stage Delay (units of τ)				
Overall DeLay				

Name _____

4. (20 points) Sizing: In the complex gate below, nmos transistor E is sized as 2W times wider than a unit transistor.





B

a) Annotate each transistor with a width that makes it "compatible" with transistor E's width of 2W.

b) _____What is the nominal ON resistance in relation to that for a unit inverter (R).

c) Draw a stick diagram of this cell Note the left-most poly is input B.

d) Annotate the stick figure above to show all capacitances. Distinguish between contacted and uncontacted capacitances. If you reach a point where two transistors of different widths have their source and drain on the same strip of diffusion, take the larger width as the capacitance,

e) _____ What is Cout (aggregate diffusion capacitance on Y)?

f) Compute Cin: A = _____;, B = _____ C = _____ D = _____ E = _____

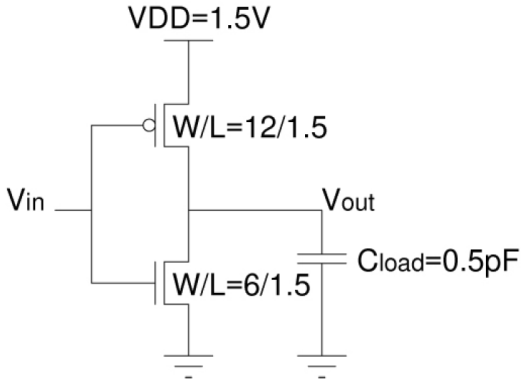
g) _____ What is the logical effort for this gate as seen by input A??

h) Draw an Elmore model and compute the parasitic delay when A=0; B=C=D=1, and E goes from 0 to 1

Name _____

5. Consider the inverter pictured below. W and L are in um.

Parameter	NMOS	PMOS
V_T (threshold)	0.5V	-0.5V
u (mobility)	200 cm^2/Vs	100 cm^2/Vs
λ (1/Vearly)	0.1 V^{-1}	0.1 V^{-1}
t_{ox} (oxide)	15nm = $1.5x10^{-6}$ cm	



a) Derive the Long Channel Parameters for both these transistors. To make the math simpler, you may assume $k_{ox}\varepsilon_0$ is 3E-13 V/cm. Make sure you are using consistent units. (Note: in book Example 2.1 the "262" should be in units of uA/V^2 not A/V^2.)

b) Set up and solve an equation that tells you when Vin drives Vout to 0.75V. What is both Vin and the current in this case? Ignore the C_{load} for now.

c) The IV curve on the next page is a partial IV curve for the above NMOS transistor. Determine what the y-axis units are for this chart, and annotate. Then draw the curve for the Vgs corresponding to the Vin value you computed in the last question. Take channel length modulation into account (Section 2.4.2 p. 78) (Remember that $\lambda = 1/V_A$, where V_A, is the "Early voltage" of page 78).

d) .Determine the effective resistance for the NMOS transistor (use the approach of Section 4.3.7 page 154).

e) Estimate the delay in ps for this gate for the case Vin goes from 0 to 1. Assume the only capacitance is the indicated load capacitance.

Name _____