EECS 151/251A Midterm 2 Review Session
Topics in Scope

- RISC-V ISA, pipelining, hazard resolution (data + control), pipeline diagrams, critical path analysis
- LUTs, how to use them to implement logic functions, how to combine them
- MOS switch model, sketching VTCs, find logical function from schematic
- MOS resistance/capacitance relationship with width
- Pass transistor gates, static CMOS gates (+ construction from logic formula), inverter VTC and switching threshold
- Inverter/Gate sizing, logical effort for minimum delay logic chain sizing
- RC delay, Elmore delay approximation
- Wire models
- Power/energy, energy on capacitor/drawn from supply, dynamic power
- Adders (ripple carry, carry bypass, carry select, CLA)
Topics NOT in Scope

- Multipliers
- Flip-flops/latches
RISC-V Datapath/Pipelining/Hazards
5. [12pts/12pts] MIPS microarchitecture.

Consider the design of the single-cycle MIPS processor, as we were discussed in class. Suppose you want to pipeline this design using 2 pipeline stages. Functionality is divided into 2 pipeline stage as shown below:

<table>
<thead>
<tr>
<th>stage 1</th>
<th>stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>instruction fetch/decode</td>
<td>ALU operation</td>
</tr>
<tr>
<td>regfile access</td>
<td>Dmem access</td>
</tr>
</tbody>
</table>

In the space below, explain what hazards will result from this pipelining (prior to any measures to remove the hazards). For each hazard, i) describe why it occurs, ii) write a short MIPS assembly language instruction sequence to demonstrate a case when the hazard would occur, and iii) describe what you can do to deal with the hazard. Write your answers in the space provided. (You might not need all answer spaces.)
(a) Hazard 1

i. Control hazard occurs on every branch instruction because the branch comparison is done in stage 2 and the branch target address is needed in stage 1.

ii. Any instruction sequence with a branch.

iii. 1) Stall the pipeline to allow the branch instruction to complete, 2) implement branch prediction.

(b) Hazard 2

i. Data hazard occurs on r-type instruction dependent on the immediately preceeding instruction, because ALU output is available sometime in stage 2 but needed at the beginning of stage 2.

ii.

\[
\begin{align*}
\text{add } &\$1, \$2, \$3 \\
\text{add } &\$4, \$5, \$1
\end{align*}
\]

iii. 1) Stall the pipeline to allow r-type to store results back to register file, 2) Selectively feed output of ALU to ALU input register.
(c) Hazard 3

i. Data hazard on load from memory when an instruction is dependent on load instruction immediately preceeding. Result from memory load is available at end of stage 2 but needed at beginning of stage 2.

ii.

\[
\begin{align*}
&\text{l}w \; $1, \; 0($1) \\
&\text{add} \; $4, \; $5, \; $1
\end{align*}
\]

iii. 1) Selectively direct output of memory to input register of ALU, 2) stall the load instruction.
FPGAs / LUTs
1. [4pts] Using 6-LUTs only (each with inputs \( x_0, x_1, \ldots, x_5 \)) and output \( y \), draw a circuit that implements an 8-LUT.
CMOS Gates
9. [4pts] Draw a CMOS transistor-level circuit for a 2-input exclusive-or function (try to minimize the number of transistors).

\[ f = A \oplus B = \overline{AB} + \overline{A}B \]

\[ f = f = \overline{A \overline{B} + \overline{A}B} = (\overline{A+B}) \cdot (\overline{A} + \overline{B}) \]
c) Implement the logic function for $C_0$ as a complex static CMOS gate (3 Pts).

$$C_0 = A_1B_1 + A_1A_0B_0 + A_0B_1B_0$$
\[ C_0 = A_1 B_1 + A_1 A_0 B_0 + A_0 B_1 B_0 \]
\[ \overline{C_0} = A_1 B_1 + A_1 A_0 B_0 + A_0 B_1 B_0 \]
\[ \overline{C_0} = A_1 B_1 \cdot A_1 A_0 B_0 \cdot A_0 B_1 B_0 \]
\[ C_0 = (A_1 + B_1) \cdot (A_1 + A_0 + \overline{B_0}) \cdot (A_0 + \overline{B_1} + \overline{B_0}) \]
\[ \overline{C_0} = (A_1 + B_1) \cdot (A_1 + A_0 + \overline{B_0}) \cdot (A_0 + \overline{B_1} + \overline{B_0}) \]
Power / Energy
(j) [1pt] The *switching energy* corresponding to a change in value on the output of a CMOS logic gate is given by $E = \frac{1}{2}CV^2$. What does “$C$” represent in the actual circuit?

(k) [1pt] Based on the equation above, if the gate is switching at a frequency of “$f$”, what is its average *power consumption*?
(j) [1pt] The *switching energy* corresponding to a change in value on the output of a CMOS logic gate is given by $E = \frac{1}{2}CV^2$. What does “C” represent in the actual circuit?

(k) [1pt] Based on the equation above, if the gate is switching at a frequency of “f”, what is its average *power consumption*?

- $C =$ total parasitic capacitance of every gate (input and output)
- Power = [Energy/Time]
- $P_{avg} = C \cdot V^2 \cdot f$
A NOR Gate (G1) with input capacitance $C_1$ is driving a wire with total resistance $R_W$ and total capacitance $C_W$, and an inverter G2 is driving an external load of $C_L$. The on-resistance of G1 and G2 are $R_1$ and $R_2$, respectively. Assume $\gamma = 1$ and $2R_N = R_P$. The power supply has a voltage of $V_{DD}$.

\[
Z_1 = R_1 (\gamma C_1 + C_W + \frac{C_2}{2}) + R_w (\frac{C_W}{2} + C_2)
\]

\[
Z_2 = R_2 (\gamma C_2 + C_L)
\]

**a)** Determine the delay between input A to Out? (4 Pts)
A NOR Gate (G1) with input capacitance $C_1$ is driving a wire with total resistance $R_W$ and total capacitance $C_W$, and an inverter with input capacitance $C_2$. Inverter G2 is driving an external load of $C_L$. The on-resistance of G1 and G2 are $R_1$ and $R_2$, respectively. Assume $\gamma = 1$ and $2R_N = R_P$. The power supply has a voltage of $V_{DD}$.

a) Determine the delay between input A to Out? (4 Pts)

$$t_1 = \ln 2 \left( R_1 \left( \frac{6}{5} C_1 + C_W + C_2 \right) \right)$$
$$t_2 = \ln 2 \left( R_2 \left( C_2 + C_L \right) \right)$$

$$t_{total} = t_1 + t_2$$
[PROBLEM 3] Energy (10 pts)

a) (4 pts) How much energy is drawn from the 1V power supply in the circuit shown below in Fig 3a when In steps from 0V to 1V? How about when In steps from 1V to 0V? You can ignore all capacitors associated with the transistors inside of the inverter.

\[ E = V_{DD} \cdot C \left[ V_{DD} - 0 \right] = C V_{DD}^2 = 25 \text{ fJ} \]
[PROBLEM 3] Energy (10 pts)

a) (4 pts) How much energy is drawn from the 1V power supply in the circuit when \( V_\text{in} \) steps from 0V to 1V? How about when \( V_\text{in} \) steps from 1V to 0V? The answer should be associated with the transistors inside of the inverter.

\[
\begin{align*}
E &= \int_{t=0}^{\infty} P \, dt = \int_{t=0}^{\infty} V_\text{DD} \cdot i_c \, dt = \int_{t=0}^{\infty} V_\text{DD} \cdot C \frac{dV_c}{dt} \, dt = \int_{V_c(0)}^{V_c(\infty)} V_\text{DD} \cdot C \left[ V_c(\infty) - V_c(0) \right] \\
E &= V_\text{DD} \cdot C \left[ V_\text{DD} - 0 \right] = CV_\text{DD}^2 = 25 \text{ fJ}
\end{align*}
\]
[PROBLEM 3] Energy (10 pts)

a) (4 pts) How much energy is drawn from the 1V power supply in the circuit shown below in Fig 3a when In steps from 0V to 1V? How about when In steps from 1V to 0V? You can ignore all capacitors associated with the transistors inside of the inverter.

\[ I_{\text{P}} \rightarrow 0V \]

\[ V_{\text{DD}} \quad V_{\text{DD}} \]

\[ I_{\text{C}} = -I_{\text{E}} \]

\[ P = V_{\text{DD}} \cdot I_{\text{E}} + V_{\text{DD}} \cdot I_{\text{C}} = 0 \]

\[ E = 0 \]
b) (6 pts) How much energy is pulled out of the 1V power supply in the circuit shown below when $I_n$ steps from 1V to 0V? How about when $I_n$ steps from 0V to 1V? You can ignore all capacitors associated with the transistors inside of the inverter.

![Circuit Diagram]

Figure 3b
b) (6 pts) How much energy is pulled out of the 1V power supply in the circuit shown below when \( I_n \) steps from 1V to 0V? How about when \( I_n \) steps from 0V to 1V? You can ignore all capacitors associated with the transistors inside of the inverter.

\[
\begin{align*}
E &= \int_{t=0}^{\infty} P \, dt = V_{DD} \cdot C_i \cdot \int_{t=0}^{\infty} dV_i = V_{DD} \cdot C_i \cdot \left[ V_i(\infty) - V_i(0) \right] = V_{DD} \cdot C_i = 15 \text{ fJ}
\end{align*}
\]
b) (6 pts) How much energy is pulled out of the 1V power supply in the circuit shown below when $I_n$ steps from 1V to 0V? How about when $I_n$ steps from 0V to 1V? You can ignore all capacitors associated with the transistors inside of the inverter.

\[ E = V_{DD} \cdot C_2 \cdot (V_2(\infty) - V_2(0)) = V_{DD}^2 \cdot C_2 = 25 \text{ fJ} \]
Logical Effort / RC Delay
a) (8 pts) Implement the function \( F = \overline{A+B+C\cdot D} \) with a complex static CMOS gate. Assuming that for this process \( R_P = 0.75*R_N \) (i.e., for the same width, a PMOS has 0.75 times the resistance of an NMOS), size your gate so that the worst-case pull up resistance is equal to the worst-case pull-down resistance.
a) (8 pts) Implement the function $F = \overline{A+B+C \cdot D}$ with a complex static CMOS gate. Assuming that for this process $R_P = 0.75*R_N$ (i.e., for the same width, a PMOS has 0.75 times the resistance of an NMOS), size your gate so that the worst-case pull up resistance is equal to the worst-case pull-down resistance.
b) **(8 pts)** What is the logical effort of this gate from the A and C inputs?

\[
LE_A = \frac{L_{gate} \cdot C_{in,\text{gate}}}{R_{in} \cdot C_{in,\text{inv}}} = \frac{13}{5}
\]

\[
LE_C = \frac{17}{7}
\]
b) \( (8 \text{ pts}) \) What is the logical effort of this gate from the A and C inputs?

Reference inverter with same \( R \) as the gate:

\[
LE_A = \frac{13/4}{7/4} = \frac{13}{7}
\]

\[
LE_C = \frac{17/4}{7/4} = \frac{17}{7}
\]
[PROBLEM 1] Logical Effort (15 pts)

Consider the following network of CMOS gates. Assume that the first inverter connected to IN is of size 1. Assume $t_{p0}$ and $\gamma$ equal to 1.

![Diagram of CMOS gates with inputs and outputs labeled](image)

a) Determine the Path Effort between IN and any of the two outputs OUT. (3 Pts)
[PROBLEM 1] Logical Effort (15 pts)

Consider the following network of CMOS gates. Assume that the first inverter connected to IN is of size 1. Assume $t_p$ and $\gamma$ equal to 1.

![Circuit Diagram]

a) Determine the Path Effort between IN and any of the two outputs OUT. (3 Pts)

$$Path \ Effort = FGIS$$
$$= (45) \left( \frac{5}{3} \right) \left( \frac{5}{3} \right) (2)$$
$$= \boxed{250}$$
b) Size $X$ and $Y$ to give the minimum delay and determine that minimum delay. (4 Pts)
Sp16 Midterm 2

\[ f_1 = f_2 = f_3 = \frac{3}{\sqrt{250}} = 6.3 \]

\[ X = 6.3 \]

\[ f_2 g_2 = 6.3 \quad \rightarrow \quad \frac{Y}{6.3} \cdot \left( \frac{5}{3} \right) = 6.3 \]

\[ Y_T = 23.8 \]

\[ Y = 11.9 \]
\[ t_p = t_{p_{0}} \left( (\gamma + f_1) + (3\gamma + f_2) + \left( \frac{10}{3} \gamma + f_3 \right) \right) \]

\[ = t_{p_{0}} \left( (1 + 6.3) + (2 + 6.3) + (3 + 6.5) \right) \]

\[ = \frac{24.9}{t_{p_{0}}} \]
a) Assume that the delay of a minimum size inverter is given by $t_p = 1+f$ (that is $t_{p0} = 1$ and $\gamma = 1$). $f$ is the fanout. For each of the configurations shown, determine the delay between In and Out. (3 points)
a) Assume that the delay of a minimum size inverter is given by $t_p = 1 + f$ (that is $t_{p0} = 1$ and $\gamma = 1$). $f$ is the fanout. For each of the configurations shown, determine the delay between In and Out. (3 points)

(a) $t_p = 1 + 64 = 65$

(b) $t_p = (1+16) + (1+8) = 22$

(c) $t_p = (1+9) + (1+3) + \left(1 + \frac{6}{3}\right) = 15.33$
b) Assuming that the input inverter sizes are fixed to 1, size the other inverters so that the delay is minimized for each of the three configurations, and determine the corresponding delay. (4 points)
Sp16 Midterm 1

(a) \( \text{same} \)

(b) \( f = \sqrt{64} = 8 \) \( \rightarrow \) \( t_F = (1+8) + (1+8) = 18 \)

(c) \( f = \sqrt[3]{64} = 4 \) \( \rightarrow \) \( t_F = 15 \)
[PROBLEM 2] Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 predecoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to 2C_cell.

\[ C_{in} = \frac{C_{cell}}{2} \]

a) (4 pts) What is the path effort from A0 to WL0?
Fa16 Midterm 2

PROBLEM 2: Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 pre-decoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to 2C_{cell}.

\[ C_{in} = 2C_{cell} \]

**Figure 2a**

\[ \text{PE} = H = F GB \]

\[ F = \frac{128 C_{cell}}{2 C_{cell}} = 64 \]

\[ L = E = G = \frac{5}{3} \cdot 1 \cdot 1 \cdot \frac{4}{3} \cdot 1 = \frac{20}{9} \]

\[ B = 2 \cdot 1 \cdot 1 \cdot 4 \cdot 1 = 8 \]

\[ \text{PE} = H = 64 \cdot \frac{20}{9} \cdot 8 = \frac{10240}{9} \approx 1137.8 \]
[PROBLEM 2] Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 predecoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to \(2C_{\text{cell}}\).

\[C_{\text{in}} = 2C_{\text{cell}}\]

\(A_0 \rightarrow \text{a} \rightarrow \text{b} \rightarrow \text{pd1}_00 \rightarrow \text{c} \rightarrow \text{d} \rightarrow \text{WL0}\)

\[C_{\text{load}} = 128C_{\text{cell}}\]

Figure2a

b) (4 pts) What EF/stage minimizes the delay of this decoder?
[PROBLEM 2] Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 predecoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to 2\(C_{out}\).

\[
\frac{E_F}{\text{Stage}} = h = \sqrt[N]{H} \quad N=5
\]

\[
h = \sqrt[5]{1137.8} \approx 4
\]
[PROBLEM 2] Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 predecoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to $2C_{cell}$.

$C_{in} = 2C_{cell}$

---

c) (4 pts) Size the gates to minimize the delay from A0 to WL0.
[PROBLEM 2] Memory Decoders (12 pts)

Shown below is the critical path of a decoder for a 16x128 SRAM array. This decoder has been implemented by using a 2-4 predecoder, followed by a 4-16 final decoder. Both true and complementary addresses, A0-A3, are available, and the input capacitance per address bit is limited to $2C_{\text{cell}}$.

Size $d$:

\[ f_d = \frac{n}{g_d} = \frac{4}{1} = 4 \]

\[ d = \frac{128 \cdot C_{\text{cell}}}{f_d} = \frac{128 \cdot C_{\text{cell}}}{4} = 32 \cdot C_{\text{cell}} \]

Size $c$:

\[ f_c = \frac{n}{g_c} = \frac{4}{3} = 3 \]

\[ c = \frac{32 \cdot C_{\text{cell}}}{3} \approx 10.7 \cdot C_{\text{cell}} \]

Size $b$:

\[ f_b = \frac{n}{g_b} = \frac{4}{1} = 4 \]

\[ b = \frac{4 \cdot \frac{32}{3} \cdot C_{\text{cell}}}{4} = \frac{32}{3} \cdot C_{\text{cell}} \approx 10.7 \cdot C_{\text{cell}} \]

Size $a$:

\[ f_a = \frac{n}{g_a} = \frac{4}{1} = 4 \]

\[ a = \frac{\frac{32}{3} \cdot C_{\text{cell}}}{4} = \frac{8}{3} \cdot C_{\text{cell}} \approx 2.7 \cdot C_{\text{cell}} \]
PROBLEM 1. Logical Effort and Gate Sizing (18 points + BONUS 6 pts)

\[ C_{\text{in}} = 1 \text{fF} \]

\[ C_{\text{load}} = 300 \text{fF} \]

a) (4 pts) What is the path effort from In to Out?

b) (2 pts) What EF/stage minimizes the delay of this chain of gates?

c) (6 pts) Size the gates to minimize the delay from In to Out.

<table>
<thead>
<tr>
<th>Size</th>
<th>Value (fF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>
PROBLEM 1. Logical Effort and Gate Sizing (18 points + BONUS 6 pts)

\[ C_{in} = 1 \text{fF} \]

\[ C_{load} = 300 \text{fF} \]

a) (4 pts) What is the path effort from In to Out?

\[ PE = \pi \cdot \frac{1}{a} \cdot \frac{1}{b} \cdot \frac{1}{c} \]

\[ = \left( \frac{6}{5}, \frac{4}{5} \right) \cdot 2 \cdot \frac{3 \text{ms}}{1 \text{fF}} = \boxed{13.33 \text{fF}} \]

b) (2 pts) What EF/stage minimizes the delay of this chain of gates?

\[ EF_{min} = \left( PE \right)^{1/4} \]

\[ = \boxed{4.04} \]

c) (6 pts) Size the gates to minimize the delay from In to Out.

\[ \frac{3 \text{ms}}{c} = 6.04 \]

\[ \frac{c}{b} \cdot \frac{5}{3} = 6.04 \]

\[ \frac{2 \cdot b}{a} = 6.04 \]

<table>
<thead>
<tr>
<th>Size</th>
<th>Value (fF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>4.34</td>
</tr>
<tr>
<td>b</td>
<td>13.47</td>
</tr>
<tr>
<td>c</td>
<td>44.65</td>
</tr>
</tbody>
</table>
[PROBLEM 3] Inverter delay (16 pts)

Consider an inverter driving a capacitive load in 0.25\(\mu\)m CMOS technology, as shown in Figure 2.a.

\[ t_{p_{inv}} = \tau_{inv} (1 + f) \]
\[ 30\text{ps} = \tau_{inv} (1 + 1) \]

\[ f = \frac{C_{in,inv}}{C_{in,inv}} \]

All transistors have minimum length and the inverter is symmetrically sized with \(W_n = 0.5\mu\text{m}\) \(W_p = 1\mu\text{m}\). In this technology, input (gate) capacitance and intrinsic (drain) capacitances are equal, \(C_g = C_d = 2\text{fF/\mu m}\).

\[ \gamma = 1 \]

a) (4pts) When this inverter drives another identical inverter of the same size, the delay is 30ps. What is the propagation delay for driving a 15fF load?
a) (4pts) When this inverter drives another identical inverter of the same size, the delay is 30ps. What is the propagation delay for driving a 15fF load?

\[ Y = \frac{C_L}{C_g} = 1 \]

\[ t_p = t_{po} \left(1 + \frac{f}{Y}\right) = t_{po} \left(1 + \frac{1}{1}\right) = 30 \text{ps} \]

\[ t_{po} = 15 \text{ps} \approx \tau_{inv} \]

\[ \frac{f}{Y} = \frac{C_L}{C_g} = \frac{15 \text{fF}}{1.5 \mu\text{m} \cdot 2\text{fF/\mu m}} = 5 \]

\[ t_p = 15 \text{ps} \cdot \left(1 + \frac{5}{1}\right) \]

\[ t_p = 90 \text{ps} \]
Adders
17. [4pts] Assume the delay though a full-adder cell (FA) is 1 from either input to either output, and the delay through a 2-input multiplexer (mux2) is 1 (from any input to output). What is the worse-case delay through an optimized 30-bit carry-select adder?
\[ t_{p, \text{inv}} = R \left( C_p + C_L \right) = \frac{R}{2} \left( p \left( 1 + \frac{C_L}{C_p} \right) \right) = R \left( p \left( 1 + \frac{C_L}{C_{\text{in}} \cdot \gamma} \right) \right) \]

\[ t_{p, \text{gate}} = T_{\text{inv}} \left( p + \frac{f}{\gamma} \right) \]

\[ \frac{C_p}{C_{\text{in}}} = \gamma \]

\[ \gamma = \frac{T_{\text{inv}} (1 + f)}{T_{\text{inv}} (1 + \frac{f}{\gamma})}, \gamma = 1 \]
- Size pullup/pulldown delays = ref inv

\[ P = \frac{C_{\text{out, gate}}}{C_{\text{out, inv}}} = \frac{1}{C_P} \]

\[ g = \frac{3}{2} \quad P = \frac{48C_{\text{in}}}{28C_{\text{in}}} \]

- \( C_{\text{G}} \)