# EECS 151/251A Homework 6

Due Monday, March 12$^{\text{th}}$, 2018

Problem 1: RC Delay and Logical Effort Basics

Take a CMOS inverter in a process where $\gamma = \frac{C_d}{C_g}$, and the PMOS effective on-resistance is equal to K times that of the NMOS (i.e. $R_p = K \cdot R_n$) for minimally sized transistors.

(a) Draw the inverter at the transistor-level and size each FET for equal pull-up and pull-down strength. Assume the NMOS is of size '1', normalized to the minimum width of the process.
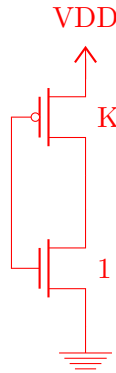
(b) Write down the delay of this unloaded inverter using RC time constants in terms of γ, R_n, and C_g. Call this delay t_p. Assume input step transitions.

$$
\begin{aligned}
t_p = \ln(2) \cdot R_{eq} \cdot C_{eq} &= \ln(2) \cdot R_n \cdot (C_{d,n} + C_{d,p}) \\
&= \ln(2) \cdot R_n \cdot (\gamma C_g + K\gamma C_g) \\
&= \ln(2) \cdot R_n \cdot (K+1)\gamma C_g
\end{aligned}
$$

(c) Now let the inverter have a load of C_L attached to it's output. Call the quantity $f = C_L/C_{in}$. Write the delay of the inverter in terms of γ, t_p, and f.

$$
\begin{aligned}
C_{in} = (W_p + W_n) \cdot C_g &= (K+1) \cdot C_g \\
t_{inv,loaded} = \ln(2) \cdot R_n \cdot (\gamma C_g + K\gamma C_g + C_L) \\
&= \ln(2) \cdot R_n \cdot (C_g \gamma (K+1) + C_L) \\
\text{Multiplying by } \frac{C_{in}}{C_{in}} : &= \ln(2) \cdot R_n (K+1)C_g \cdot (\gamma + \frac{C_L}{C_{in}}) \\
&= \frac{t_p}{\gamma} \cdot (\gamma + f)
\end{aligned}
$$

We notice that the delay isn't affected by K, as expected. We also see a linear relation between inverter loading and delay.

No matter how we large we size the inverter, its intrinsic delay is always the same, but a larger inverter is able to drive a large external load faster, at the expense of increasing its input capacitance.

In the lecture slides we call the quantity $\frac{t_p}{\gamma}$ as t_{inv} which is treated as a process constant.

(d) Recall that the gate delay for any CMOS gate can be expressed as:

$$t_{p,gate} = t_{inv}(p + f \cdot LE)$$

where t_{inv} is a process constant, p is the intrinsic delay of the gate, and LE is the logical effort of the gate.

What are these values for an inverter?

$$\text{Comparing the equation from above: } t_{inv} = \frac{t_p}{\gamma}$$
$$p_{inv} = \gamma$$
$$LE_{inv} = 1$$

(e) Take the following sizing of a NAND2 gate. What is W_p for inverter equivalent delay? Prove that $W_{n,a} = W_{n,b} = 2$ so the gate has equivalent self-loaded delay to an inverter and consumes minimal area.
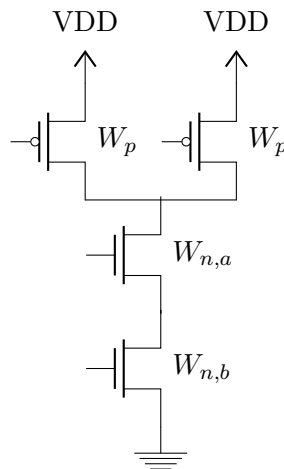
$$W_p = K \text{ for equivalent inverter pull-up strength}$$
$$\text{Area } = W_{n,a} + W_{n,b}$$
$$R_{n,a} + R_{n,b} = R_{inv} \rightarrow \frac{R_n}{W_{n,a}} + \frac{R_n}{W_{n,b}} = R_n$$
$$\text{This relation must hold: } W_{n,a} + W_{n,b} = W_{n,a} \cdot W_{n,b}$$

We can then show that to minimize area under the constraint, $W_{n,a} = W_{n,b} = 2$.

(f) Find the logical effort of this NAND2 gate for both inputs. What happens to the LE as K increases, and why does it make intuitive sense?

The LE of a gate sized such that it has inverter-equivalent pull-up and pull-down strength is:

$$LE_{gate} = \frac{C_{in,gate}}{C_{in,inv}}$$

$$LE_{nand2} = \frac{KC_g + 2C_g}{KC_g + C_g} = \frac{K+2}{K+1}$$

We notice that as K increases, the LE tends towards 1. This can be intuitively understood as the minimal inverter's PMOS needing ever increasing upsizing, which decreases the relative difference between the inverter and NAND2 input capacitance.

(g) What is the intrinsic delay of a NAND2 gate (i.e. find p_{NAND2})? Make the assumption that $K = 2$.

$$t_{NAND,RC,loaded} = t_{NAND,gate,delay,equ}$$

$$\ln(2) \cdot R_n \cdot 4 \cdot C_g(\gamma + f) = \ln(2) \cdot R_n \cdot 3 \cdot C_g(p_{NAND2} + \frac{4}{3}f)$$

$$4(\gamma + f) = 3(p_{NAND2} + \frac{4}{3}f)$$

$$p_{NAND2} = \frac{4\gamma}{3}$$

Problem 2: Gate Delay Optimization and Branching

Size the gates in the chain below for minimum delay. As usual, all gates are sized so that the pull-up and pull-down resistance match that of the reference inverter.

First, calculate the logical effort for each complex gate: $LE_a = 4/3, LE_b = 5/3, LE_d = 5/3, LE_e = 7/3$.
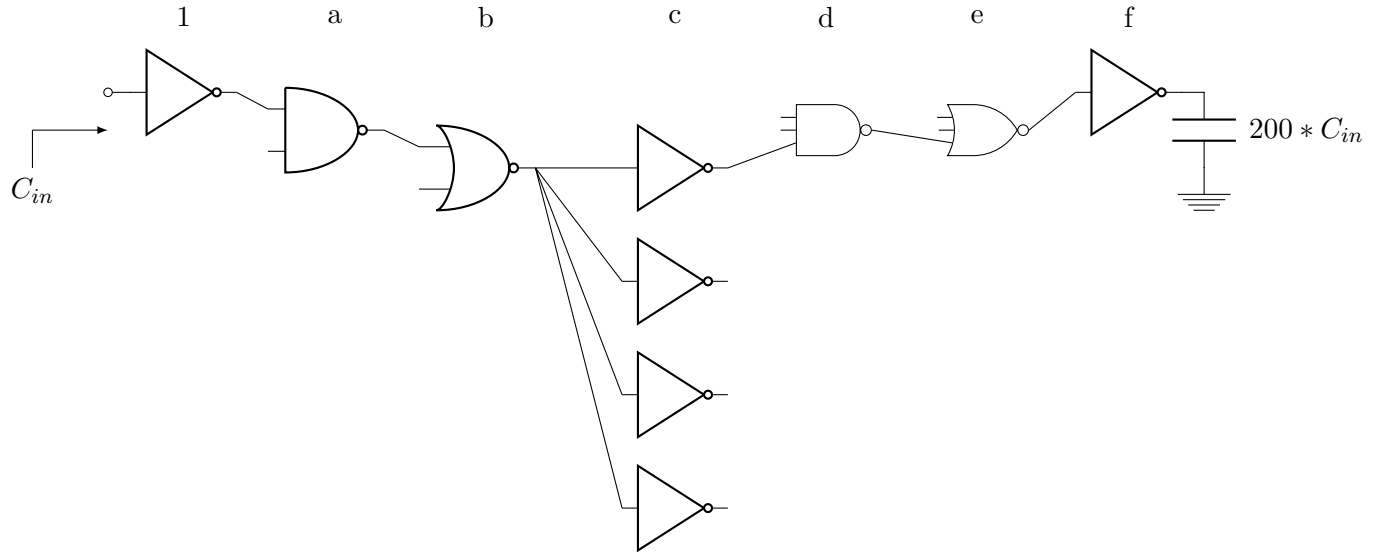
The path effort is:

$$PE = F \cdot \prod b_i \prod LE_i = 200 \cdot 4 \cdot (4/3 \cdot 5/3 \cdot 5/3 \cdot 7/3) \approx 6914$$

The effective fanout per stage is optimally:

$$EF = \sqrt[N]{PE} = \sqrt[7]{6914} \approx 3.54$$

Therefore, working backwards from gate (f) to gate (a):

$$LE_{inv} \cdot \frac{C_L}{f} = EF \rightarrow f = \frac{C_L}{EF} = 56.5C_{in} \text{ where 'f' is the inverter size}$$

$$LE_e \cdot \frac{f}{e} = EF \rightarrow e = \frac{f}{EF} \cdot LE_e = 37.2C_{in}$$

$$LE_d \cdot \frac{e}{d} = EF \rightarrow d = \frac{e}{EF} \cdot LE_d = 17.52C_{in}$$

$$LE_{inv} \cdot \frac{d}{c} = EF \rightarrow c = \frac{d}{EF} \cdot LE_c = 4.95C_{in}$$

$$LE_b \cdot \frac{c}{b} = EF \rightarrow b = b_b \cdot \frac{c}{EF} \cdot LE_b = 9.33C_{in}$$

$$LE_{inv} \cdot \frac{b}{a} = EF \rightarrow a = \frac{b}{EF} \cdot LE_a = 3.52C_{in}$$

Problem 3: Voltage Scaling

To gain intuition about voltage scaling, it is important to estimate delay and energy as a function of voltage. The Alpha Power Law model of a transistor's current can estimate delay over a wide voltage range. According to that model:

$$I_{ON} \propto (V_{DD} - V_{TH})^{\alpha}$$

$$I_{OFF} \propto e^{\frac{-V_{TH}}{kT/q}}$$

Use this model with $\alpha = 1.3$, $V_{TH} = 0.3V$ and $V_{DD,nom} = 1V$ to answer the following questions about voltage scaling. Assume that at nominal supply voltage a given design runs at $1GHz$ and that the total capacitance is $C_{TOT} = 1nF$.

a) For $0.35V < V_{DD} < 1V$, considering dynamic energy only and an activity factor of 0.1, generate the power vs. delay (t_d) plot.

The current equation is $I_{ON} = \beta(V_{DD} - 0.3)^{1.3}$, where β is some constant (note that we keep all units in SI and remove them from the equations for simplicity). For worst-case delay we have:

$$t_d = C\frac{\Delta V}{I_{ON}} = C\frac{V_{DD}/2}{\beta(V_{DD} - 0.3)^{1.3}}$$

Another way of getting this is by using $R_{ON} = \frac{1}{2ln(2)}\frac{V_{DD}}{I_{ON}}$:

$$t_d = ln2 \cdot R_{ON}C = ln2 \cdot \frac{1}{2ln(2)}\frac{V_{DD}}{\beta(V_{DD} - 0.3)^{1.3}} \cdot C$$

Setting $t_d = \frac{1}{f_{clk}} = \frac{1}{1GHz}$ for $V_{DD} = 1$ we can get an approximate value for $\frac{C}{\beta} = 1.26 \cdot 10^{-9}$. Therefore now we have t_d as a function of V_{DD}.

The dynamic power is $P = a_{0\to1}C_{TOT}V_{DD}^2\frac{1}{t_d}$. Therefore, P as a function of V_{DD} would be:

$$P = a_{0\to1}C_{TOT}V_{DD}^2 \cdot \frac{\beta(V_{DD} - 0.3)^{1.3}}{CV_{DD}/2} = 2a_{0\to1}C_{TOT}V_{DD}\frac{(V_{DD} - 0.3)^{1.3}}{\frac{C}{\beta}}$$

Using Python, we get the power vs. delay plot on the next page.
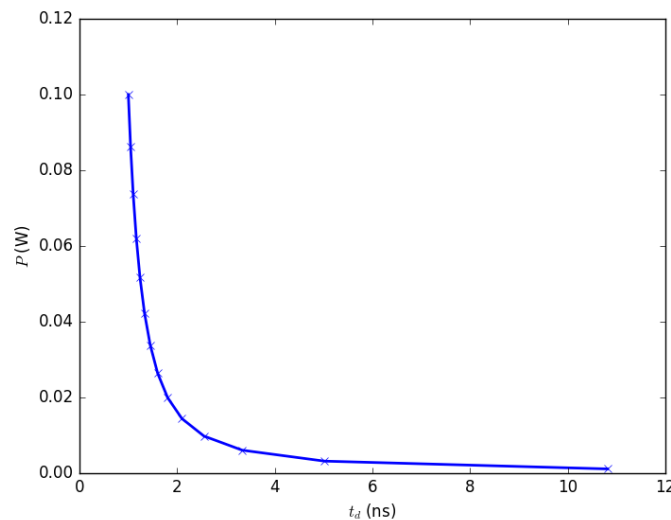
b) Now consider power as a combination of dynamic energy and leakage power. Using $I_{OFF} = 100mA$ for your given design, plot V_{DD} vs. the energy per cycle for activity factors $a = 0.1, 0.2, 0.3$. Is voltage scaling more or less effective for higher activity factors? Explain why.

The delay equation is the same as before. For power $P = a_{0\to1}C_{TOT}V_{DD}^2\frac{1}{t_d} + V_{DD}I_{OFF}$. The energy per cycle is $P \cdot t_d$ and is plotted below:

Without considering leakage, you always want to scale voltage. But if there is leakage (there always is), then energy will have a minimum. This is because at low voltage near the threshold, delay increases pretty dramatically. This means that each cycle becomes very long, and integrates a lot of leakage current (even if the leakage current is low). So the amount of leakage energy per cycle increases, and actually begins to pass the amount of dynamic energy. The higher the activity factor, the less of an impact this has, because leakage is a smaller proportion of overall energy.
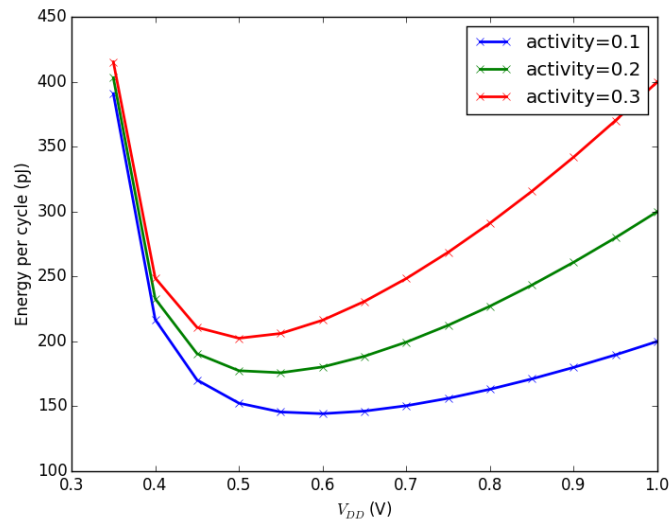
Problem 4: Elmore Delay

For the following problem, $C_G = C_D = 2fF/um$, the minimum sized (labeled as 1x in the picture) inverter has $L = 0.1um$, $W_p = 2um$, $W_n = 1um$ and for this technology $R_{n,on} = 10k\Omega/sq.$ (i.e. the resistance of an NMOS with width W and length L is equal to $10k\Omega\frac{L}{W}$) and $R_{p,on} = 20k\Omega/sq.$ (i.e. the resistance of a PMOS with width W and length L is equal to $20k\Omega\frac{L}{W}$). Note that a 6x inverter has 6 times the width of a 1x inverter

For the wire, $R_{wire} = 0.1\Omega/sq.$, the parallel plate capacitance is $C_{pp} = 20aF/um^2$ and the fringing capacitance per each side of wire is $C_{fr} = 14aF/um$. The wire widths and lengths are shown in the picture.
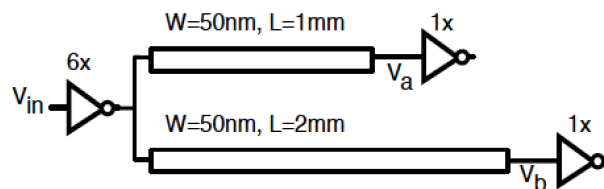
a) Using the π wire model, draw the equivalent RC switch model. What is the propagation delay from a step at Vin to Va and Vb?

- For the 1x inverter:

$$R_{N,1x} = R_{P,1x} = 10k\Omega\frac{0.1\mu m}{1\mu m} = 1k\Omega$$

$$C_{i,1x} = C_{o,1x} = (2\mu m + 1\mu m)2fF/\mu m = 6fF$$

- For the 6x inverter:

$$R_{N,6x} = R_{P.6x} = R_N/6 \approx 167\Omega$$

$$C_{o,6x} = 6 \cdot C_{o,1x} = 36fF$$

- For the top wire:

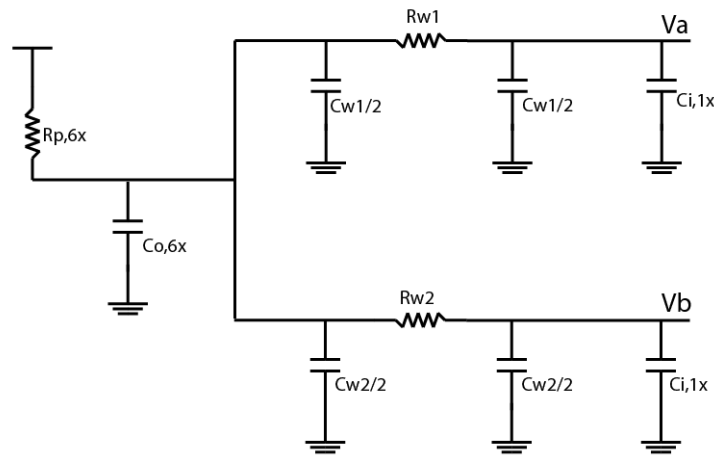$$R_{w1} = 0.1\Omega \frac{1000\mu m}{0.05\mu m} = 2k\Omega$$

$$C_{w1} = 20aF/\mu m^2 \cdot 0.05\mu m \cdot 1000\mu m + 2 \cdot 14aF/\mu m \cdot 1000\mu m = 29fF$$

- For the bottom wire:

$$R_{w2} = 2 \cdot R_{w1} = 4k\Omega$$

$$C_{w2} = 20aF/\mu m^2 \cdot 0.05\mu m \cdot 2000\mu m + 2 \cdot 14aF/\mu m \cdot 2000\mu m = 58fF$$

The equivalent RC circuit is:

The delay from input to V_a is:

$$t_{p,a} = ln2 \cdot R_{w1} \cdot (C_{i,1x} + \frac{C_{w1}}{2}) + ln2 \cdot R_{p,6x} \cdot (C_{i,1x} + \frac{C_{w1}}{2} + C_{i,1x} + \frac{C_{w2}}{2} + \frac{C_{w1}}{2} + \frac{C_{w2}}{2} + C_{o,6x})$$

$$\Rightarrow t_{p,a} = ln2 \cdot 62.38ps = 43.24ps$$

The delay from input to V_b is:

$$t_{p,b} = ln2 \cdot R_{w2} \cdot (C_{i,1x} + \frac{C_{w2}}{2}) + ln2 \cdot R_{p,6x} \cdot (C_{i,1x} + \frac{C_{w2}}{2} + C_{i,1x} + \frac{C_{w1}}{2} + \frac{C_{w2}}{2} + \frac{C_{w1}}{2} + C_{o,6x})$$

$$\Rightarrow t_{p,b} = ln2 \cdot 162.55ps = 112.27ps$$

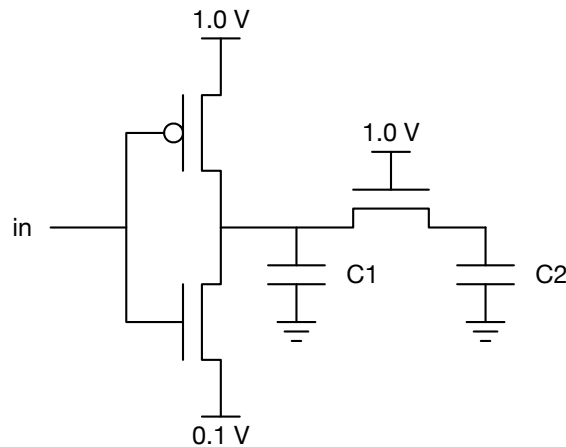b) What is the skew (difference in arrival time between Va and Vb)?

$$t_{skew} = |t_{p,b} - t_{p,a}| = 69.03ps$$

Problem 5: Energy

Given the circuit below, answer the following questions. Assume that the capacitances associated with the FETs are negligible.

$C_1 = C_2 = 90 fF$

$V_{th,n} = |V_{th,p}| = 0.4V$

Part a

How much energy is stored in each capacitor after a 1V to 0V transition on `in`? Please show your work.

$E_1 = \frac{1}{2} \cdot C_1 \cdot V_1^2 = \frac{1}{2} \cdot 90 fF \cdot (V_{DD})^2 = \frac{1}{2} \cdot 90 fF \cdot (1.0V)^2 = 45.0 fJ$

$E_2 = \frac{1}{2} \cdot C_2 \cdot V_2^2 = \frac{1}{2} \cdot 90 fF \cdot (V_{DD} - V_{th,n})^2 = \frac{1}{2} \cdot 90 fF \cdot (0.6V)^2 = 16.2 fJ$

Part b

How much energy was drawn by the supply in part a? Is this what you expected? Please show your work.

$E_{supply,1->0} = \int_0^\infty i_{DD}(t) \cdot V_{DD} \; dt = V_{DD} \int_0^\infty C_1 \frac{dV_1}{dt} + C_2 \frac{dV_2}{dt} \; dt = V_{DD} \cdot \left(C_1 \int_0^\infty dV_1 + C_2 \int_0^\infty dV_2 \right)$

$E_{supply,1->0} = V_{DD} \cdot C_1 \cdot (V_1(\infty) - V_1(0)) + V_{DD} \cdot C_2 \cdot (V_2(\infty) - V_2(0))$

$E_{supply,1->0} = 1.0V \cdot 90 fF \cdot (1.0V - 0.1V) + 1.0V \cdot 90 fF \cdot (0.6V - 0.1V) = 126 fJ$

Part c

How much total energy is dissipated when the input goes from 1V to 0V and then back to 1V? Please show your work.

We can use the answer from part b as the energy drawn from the supply from a 1V to 0V transition.

Then we can calculate the amount of energy "returned" during a 0V to 1V transition.

$$E_{supply,0->1} = \int_0^\infty i_{SS}(t) \cdot 0.1V \; dt = 0.1V \int_0^\infty C_1 \frac{dV_1}{dt} + C_2 \frac{dV_2}{dt} \; dt = 0.1V \cdot \left(C_1 \int_0^\infty dV_1 + C_2 \int_0^\infty dV_2 \right)$$

$$E_{supply,0->1} = 0.1V \cdot C_1 \cdot (V_1(\infty) - V_1(0)) + 0.1V \cdot C_2 \cdot (V_2(\infty) - V_2(0))$$

$$E_{supply,0->1} = 0.1V \cdot 90fF \cdot (0.1V - 1.0V) + 0.1V \cdot 90fF \cdot (0.1V - 0.6V) = -12.6fJ$$

$$E_{total} = E_{supply,1->0} + E_{supply,0->1} = 126fJ - 12.6fJ = 113.4fJ$$

We can verify that this makes sense with a quick conservation of energy test:

$$E_{supply,1->0} = \Delta U_{1->0} + E_{diss,1->0}$$

$$\Delta U_{1->0} = 90fF \cdot \frac{(1.0V)^2}{2} - 90fF \cdot \frac{(0.1V)^2}{2} + 90fF \cdot \frac{(0.6V)^2}{2} - 90fF \cdot \frac{(0.1V)^2}{2} = 60.3fJ$$

$$E_{diss,1->0} = E_{supply,1->0} - \Delta U_{1->0} = 126fJ - 60.3fJ = 65.7fJ$$

$$E_{supply,0->1} = \Delta U_{0->1} + E_{diss,0->1}$$

$$\Delta U_{0->1} = 90fF \cdot \frac{(0.1V)^2}{2} - 90fF \cdot \frac{(1.0V)^2}{2} + 90fF \cdot \frac{(0.1V)^2}{2} - 90fF \cdot \frac{(0.6V)^2}{2} = -60.3fJ$$

$$E_{diss,0->1} = E_{supply,0->1} - \Delta U_{0->1} = -12.6fJ + 60.3fJ = 47.7fJ$$

$$E_{total} = E_{diss,1->0} + E_{diss,0->1} = 65.7fJ + 47.7fJ = 113.4fJ$$

Problem 6: Processor Datapath

You are to add a *store with postincrement* instruction to a single-stage MIPS processor. The instruction **swinc** updates the index register to point to the next memory word after completing the store. **swinc $rt, imm($rs)** is equivalent to the following two instructions:

```
sw          $rt, imm($rs)
addi        $rs, $rs, 4
```

(Reminder: The first line syntax means store the word from register $rt to address imm + contents of register $rs.)

How would you modify the following datapath to accommodate this instruction? Try to add as little hardware as possible.
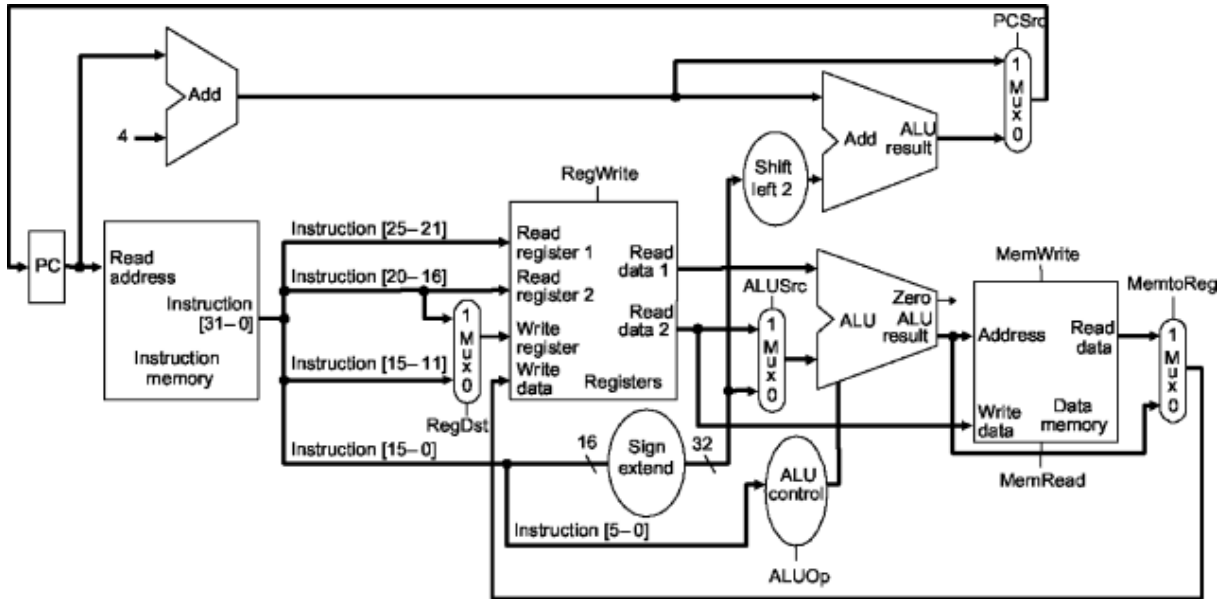
Since the main ALU is busy with calculating the store address (imm + $rs), it can't be used for adding 4 to $rs. However, the adder that normally computes the jump address is free during this instruction. Therefore, we can reuse that for this purpose. Its result will be an additional input to the MemtoReg MUX, which will be written back to $rs register.

Another solution is to use a seperate adder as in next page. Any other correct solution will also receive full credit.