CENG 4480 Final (Fall 2020)

Name: Ng Chi Hon
ID: 1155116317

Q1 (30 marks, 1 mark for each single-choice question, and 2 marks for each multiple-choice question) Short answer questions: **Circle** the right answer.

1. A circuit where the output signal power is greater than the input signal power is called amplifier / attenuator.
2. Capacitance is a measure of how much Current / Charge / Resistance a capacitor can hold.
3. Magnitude of $\frac{1}{1+ja}$ is $\frac{1}{\sqrt{1-a^2}}$ / $\frac{1}{\sqrt{1+a^2}}$ / $\frac{1}{\sqrt{1+a}}$.
4. When the Frequency of the circuit is low, the circuit is Open / Short.
5. The gain of a system with input V_{in} and output V_{out} in decibel is $20 \cdot \log_{10} \frac{V_{in}}{V_{out}}$ / $20 \cdot \log_{10} \frac{V_{out}}{V_{in}}$ / $10 \cdot \log_{10} \frac{V_{out}}{V_{in}}$.
6. Weighted Adder / R-2R DAC is impossible to fabricate a wide range of resistor values in the same IC chip.
7. **(Multiple Choices)** Comparison: Integrating / Tracking / Successive Approximation / Flash ADC is slow.
8. Sample-and-Hold Amplifier is motivated when a Slow / Fast ADC is used to sample a rapid changing signal only a short sampling point can be analyzed.
9. Successive Approximation replaces "Control Logic" by "Up-down Counter" / performs Binary search to determine the output bits
10. Compared to single-element strain gauge, four-element strain gauge is NOT sensitive to Temperature / Strain change.
11. **(Multiple Choices)** The pros of Flash ADC includes (1) Very fast for high quality audio and video / (2) Sample and hold circuit NOT required / (3) Very Low-Cost for wide bits conversion.
12. As complementary mechanisms, Accelerometer / Gyroscope has higher frequency than the other.
13. **(Multiple Choices)** Pulse Width Modulation provides Digital / Analog results with Analog / Digital means.
14. **(Multiple Choices)** In PID control, increasing K_p can Increase / Decrease the rise time; while increasing K_d can Increase / Decrease the overshoot problem.
15. The product of two Gaussian distributions is Gaussian / Binomial distribution.
16. Interface between the main memory and the processor is called Memory / Bus controller.
17. The interface, between the main memory and the processor, applies Gray Codes / Hamming Codes to address.
18. Temporal / Spatial Locality: If a particular storage location is referenced at a particular time, then it is likely that nearby memory locations will be referenced in the near future.
19. In Write-Through / Write-Back scheme, cache and main memory are updated at the same time.
20. The size of a single cache-line is 64 Bits / Bytes / KB.

21. DRAM / Phase-Change Memory / STT-RAM / Flash is volatile.
22. SRAM is faster / slower than DRAM, and larger / smaller than DRAM in cell size.
23. SRAM / DRAM requires periodic refresh.
24. NAND / NOR ROM works faster.
25. Setup Time is the time the input data must be stable after / before the clock transition occurs.
26. Time Margin measures the slack, or excess time, remaining in each clock cycle: (1) depend on only time delay of logic paths / (2) to Protect your circuit against signal cross-talk, miscalculation of logic delays, and later minor changes in the layout

Q2 (10 marks) Short answer questions: **Write** the right answer.

1. N-bit physical addresses can address 2^N memory location.
2. To reduce the glitches of ADC, low-pass filter is induced (Name one approach is sufficient).
3. In Kalman filter, we assume all process and measurement noises are with mean zero.
4. Identify all levels of the memory hierarchy in modern computer systems. Also analyze the properties of each hierarchy level, in terms of capacitance, performance and costs.

Four level low to high: On chip register, L_1 cache, L_2 Cache, Main memory, second storage, Tertiary storage.

The capacitance and cost is decrease, and performance is increase from low level to high level

$V_{out} =$

Q3 (8 marks) About Digital-to-Analog Conversion (DAC)

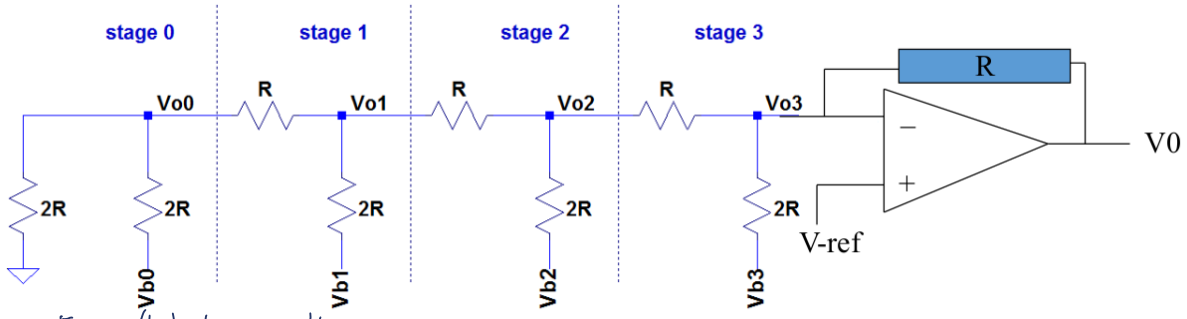
(a) Given n-bit input codes $(B_nB_{n-1}...B_1)_2$ for DAC, hence V_out can be expressed as follow. (We define High Reference Voltage as V_{+ref} and Low Reference Voltage as V_{-ref})

$$V_out = (2^nB_n + 2^{n-1}B_{n-1} + ... + B_1) \cdot \triangle V + V_{-ref} \quad (1)$$

Express how to determine $\triangle V$ using N, V_{+ref} and V_{-ref}.

$\triangle V = \frac{V_{+ref} - V_{-ref}}{2^N}$ where N is no of bit of input code

(b) For the 4-bit R-2R DAC (see the following figure), calculate V_0 in terms of $V_{b,0} - V_{b,4}$ if V_{ref} is grounded.



For 4 bit, $n = 4$

$$V_{03} = \frac{V_{b0}}{2^4R} + \frac{V_{b1}}{2^3R} + \frac{V_{b2}}{2^2R} + \frac{V_{b1}}{2^1R}$$

$$V_{out} = -V_{03}\left(\frac{R_f}{R}\right)$$

$$= -\left(\frac{V_{b0}}{16R} + \frac{V_{b1}}{8R} + \frac{V_{b2}}{4R} + \frac{V_{b1}}{2R}\right)\left(\frac{R}{R}\right)$$

$$= -\frac{V_{b0} + 2V_{b1} + 4V_{b2} + 8V_{b1}}{16R} \;//$$

Q4 (12 marks) Given the Memory Controller as shown in Figure 1, answer the following questions.

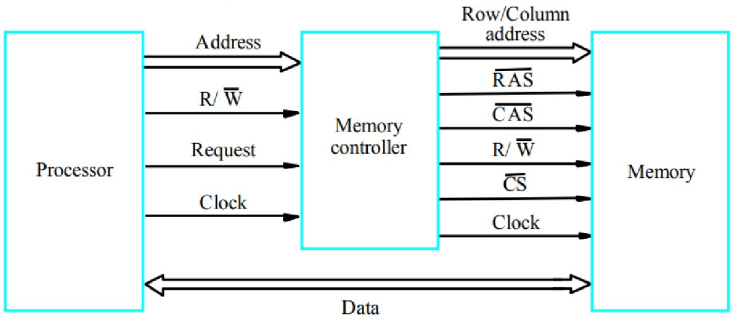


Figure 1: Memory Controllers as the interface between Processors and Memory.

(a) Reason about why DRAM requires the Memory Controllers, but SRAM may not demand them.

(b) What are $\overline{RAS}$ and $\overline{CAS}$? What are the usages of them (in particular, where are the differences between them and Row/Column Address)

(c) What are $\overline{CS}$? What are the usages of them (in particular, where are the differences between them and respective addresses)

(d) Consider a memory controller handles Non-volatile Memory, identify key changes and the reasons from a memory controller for DRAM.

save by ICIT, if no refreshing, capictor lose energy and memory will forgot their state

a) Since DRAM data is dynamic memory that each memory need to refresh many time in every second. A memory controller need keep refreshing. Also, DRAM have slightly more complex interface that have time-multiplex signal to reduce pin no.

SRAM no need refresh and less size allow simple read write operation from CPU, So need need memory controller.

b) $\overline{RAS}$ is Row Access strobe & $\overline{CAS}$ is column Access strobe.
For DRAM, each address has 64-bits and to read/write the memory, CPU will send 64 bit address to memory controller. The memory controller will convert address to row address and column address.
Then CPU will activite RAS line to specifly the row address where data be find (High bit).
And after that CPU activite CAS line to specifly the colum address where data be find (low bit)
and data is located and output by data line

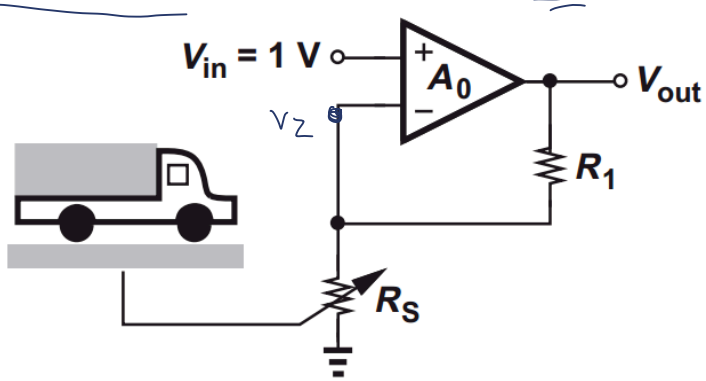
RAS & CAS is signal to activate and map to particulaur location of memory
but Row / column address is just the location of memory.

c) $\overline{CS}$ is chip-select signal. In memory there are many chip that each chip have a no. of Rank which is memory space. $\overline{CS}$ can send the signal to the chip that address belong to. respective address is just the location of data

d) The memory controller in SRAM don't need to handle refresh in SRAM,

$\beta W + R_0$ R_S

Q6 (8 marks) A truck weighing station incorporates a sensor whose resistance varies linearly with the weight: $R_S = \beta W + R_0$. Here R_0 is a constant value, β is a proportionality factor, and W the weight of each truck. Suppose R_S plays the role of R_2 in the noninverting amplifier (see the following figure). Also, $V_{in} = 2$ V.
Determine the gain of the system, defined as the change in V_{out} divided by the change in W.



$R_S = \beta W + R_0$ Let V_2 be voltage input in -ve of amplifier

$$V_2 = \frac{R_S}{R_S + R_1} \cdot V_{out}$$

$$\therefore V_{out} = V_2 \cdot \frac{R_S + R_1}{R_S} = V_2 \cdot \left(1 + \frac{R_1}{\beta W + R_0}\right)$$

For ideal summing point, $V_2 = V_{in} = 2V$

By question defination, voltage Gain is $\frac{V_{out}}{W}$

$$\therefore \text{Voltage Gain} = \frac{V_{out}}{W}$$

$$= \frac{V_2}{W} \cdot \left(1 + \frac{R_1}{R_S}\right)$$

$$= \frac{2}{W} + \frac{2R_1}{\beta W^2 + R_0 W}$$

Q6 (8 marks) The general equation of a liner estimate system is like $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{w}_{t+1}$. Consider the following three different systems:

(a) $x(t) = 1.32x(t-1) + \omega_t$

(b) $x(t) = 2.32x(t-1) - 0.76x(t-2) + \omega_t$

(c) $x(t) = 2.32x(t-1) - 0.76x(t-2) + 0.91x(t-3) + \omega_t$

Kalman Filter is used to estimate $x(t)$ (here $x(t)$ is a scalar). Please give the formulations of state transition matrix $\mathbf{A}$ and noise vector $\mathbf{w}_t$.

6)

a)

$x(t) = 1.32x(t-1) + 0\,x(t-2) + wt$

$x(t-1) = 0\,x(t-2) + 1\,x(t-1) + 0$

$$\begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1.32 \end{bmatrix} \begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ wt \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1.32 \end{bmatrix} \quad w_t = \begin{bmatrix} 0 \\ wt \end{bmatrix}$$

b)

$$\begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -0.76 & 2.32 \end{bmatrix} \begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ wt \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 \\ -0.76 & 2.32 \end{bmatrix} \quad w_t = \begin{bmatrix} 0 \\ wt \end{bmatrix}$$

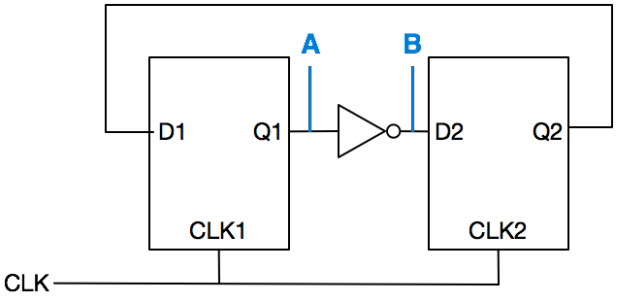
c)

$$\begin{bmatrix} x(t-2) \\ x(t-1) \\ x(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.91 & -0.76 & 2.32 \end{bmatrix} \begin{bmatrix} x(t-3) \\ x(t-2) \\ x(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ wt \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.91 & -0.76 & 2.32 \end{bmatrix} \quad w_t = \begin{bmatrix} 0 \\ 0 \\ wt \end{bmatrix}$$

Q7 (12 marks) A digital clock is important in circuit design. Please answer the following **three** questions.

(a) Given the following circuit, CLK1 = CLK2 = 25MHz; Tff = 6ns; Tsetup = 6ns. For each gate the gate delay TG = 8ns. Please calculate the time margin. Note: Tff= delay of a flip flop, Tsetup=setup time of a flip flop, and TG is delay of a gate.

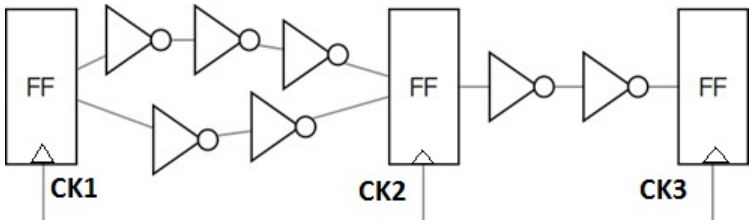


(b) In the above circuit, currently there is already one delay gate with delay TG. How many more similar delay gates can you insert between A and B without creating error?

(c) Sometimes we can take advantage of clock skew. For example, given the following circuit, TG=8ns, Tff=6ns, Tsetup=6ns, a master clock TCLK is driving CK1 and CK2 and CK3.

- The delay from TCLK to the clock input CK1 of the flip flop (FF) is Tc1.
- The delay from TCLK to the clock input CK2 of the flip flop (FF) is Tc2.
- The delay from TCLK to the clock input CK3 of the flip flop (FF) is 2ns.

Calculate the minimal clock period of the clock TCLK?



a) $\text{Time}_{\text{margin}} = \frac{1}{25 \cdot 10^6} - T_{ff} - T_G - T_{setup}$

$= 40\,ns - 6\,ns - 8\,ns - 6\,ns$

$= 20\,ns$ //

b) For each T_G is 8ns, maximum delay gate I can insert

$= \text{floor}(20/8) = 2$ //

c) Let FF_1 be the flipflop with CK1, FF_2 beflipflop with Ck2, FF_3 be filpflop with Ck_3

FF_1 to FF_2 : $T_{C1} + T_{ff} + 3T_G + T_{setup} \le T_{C2} + T_{CLK}$

FF_2 to FF_3 : $T_{C2} + T_{ff} + 2T_G + T_{setup} \le T_{C3} + T_{CLK}$

∴ $\begin{cases} T_{C1} + 36 - T_{C2} \le T_{CLK} \\ T_{C2} + 28 - T_{C3} \le T_{CLK} \end{cases}$

∴ minimum $T_{CLK} = \max(T_{C1} + 36 - T_2,\ T_{C2} + 26)$

To Minimize T_{CLK}, we should minimize $T_{C1} + 36 - T_2$ & $T_{C2} + 26$

∴ $T_{C1} = 0$ & $T_{C2} = 5$

Then minimize $T_{CLK} = 31ns$ //

Q8 (12 marks) Follow the Guidelines and apply your knowledge from Memory Lectures to discuss about the following **three** questions.

(a) There are growing interests in Approximate Memory, which introduces tolerable errors for data storage. **Reason about one benefit and one downside of Approximate SRAM and DRAM.**

Guidelines: 1) Approximate Memory can potentially utilize lower-than-standard timing parameters, since they can have errors; 2) For Approximate SRAM, one line of works apply varied supply voltages; 3) For Approximate DRAM, one line of works adapt lower refresh rates; 4) CAN Approximate Memory be used for all applications; and 5) HOW Software/Programmers take advantage of Approximate SRAM/DRAM

(b) DRAM access latency has been a major bottleneck in modern computer systems. One promising solution is to build in-DRAM cache to mitigate such issues. An example implementation of in-DRAM cache leverages shorter access time in near-Sense-Amplifier regions, compared with far-Sense-Amplifier regions.

Perform a critical analysis, by naming one strength and one weakness about this type of in-DRAM caches.

Guidelines: 1) WHY and HOW using a small fraction of memory regions can satisfy the needs of computations; 2) WHY and HOW the access latency can be reduced, given the fact that in-DRAM caches occupy a small fraction of memory regions; 3) HOW the manufactures can benefit or suffer from the designs of in-DRAM caches; and 4) HOW existing Operating Systems support the designs of in-DRAM caches.

(b) Modern main memory architectures are composed of DRAM. However, emerging non-volatile memory technologies (e.g. PCM) are considered as alternatives of DRAM. One outstanding example is recently-released Optane by Intel. **Use your knowledge of memory hierarchy and DRAM characteristics to reason: Why non-volatile memory technologies are considered as suitable alternatives of DRAM in general?**

Guidelines: Think about performance, capacitance and energy consumption.

A) For Approximate SRAM, it allow SRAM to have larger size of capacity which allow larger cache to assist computer working. Then the computer can work faster with large cache demand. However, Approximate SRAM utilize lower-than-standard timing parameters so the speed with low cache usage is faster than Approximate SRAM. It lose the speed advantage from SRAM to DRAM
For Approximate DRAM, it work with lower refresh rate which mean it save the refresh time and the speed and energy consumption in refresh is better. However since only one line of works adapt lower refresh rates, with high DRAM demand which use up size in DRAM, the speed will back the original and the refresh rate different make the burden on memory controller is high.

B) For in-DRAM cache, it allow DRAM to save more frequently use data in cache to increase speed of read/write operation and provides data locality for the pre-fetched pages. So it reduce long latency of DRAM compare to cache.

However, for system to make use the cache in DRAM, the memory controller need to handle which data sent to cache in DRAM and which sent to DRAM directly. It require modifications to the current computing system where some system may not support. So the in-DRAM cache compatibility is low

C)
Nonvolatile memory which uses transistor or other technology that wont lose data when power of device is shut off, to store data and DRAM use capacitor to store data
Then DRAM need to keep refreshing many times in every second to hold the data and Nonvolatile memory no need to refreshing. Nonvolatile memory save the refreshing time in main memory read/write operation. Then the performance of Nonvolatile memory is faster than DRAM in main memory operation.

Also, with DRAM keep refreshing many times, the energy consumption in long term is much larger than Nonvolatile memory.

Finally, the DRAM in main memory usually is not large due to the usual read/write operation in main memory is small and frequent. Then the size advantage with DRAM compare to Nonvolatile memory is loss.

So with advantage in performance and energy consumption, non-volatile memory technologies are considered as suitable alternatives of DRAM in general.