1 Emerging Memory Technologies [40 points]

Researchers at Lindtel developed a new memory technology, L-RAM, which is non-volatile. The access latency of L-RAM is close to that of DRAM while it provides higher density compared to the latest DRAM technologies. L-RAM has one shortcoming, however: it has limited endurance, i.e., a memory cell stops functioning after 10^6 writes are performed to the cell (known as cell wear-out).

(a) [15 points] Lindtel markets a new computer system with L-RAM to have a lifetime of 2 years and the following specifications:

- 4 GBs of L-RAM as main memory with a *perfect* wear-leveling mechanism, i.e., writes are equally distributed over all the cells of L-RAM.
- The processor is in-order and there is no memory-level parallelism.
- It takes 4 ns to send a memory request from the processor to the memory controller and it takes 20 ns to send the request from the memory controller to L-RAM. The write latency of L-RAM is 40 ns.
- L-RAM is word-addressable. Thus, each write request writes 8 bytes to memory.

A student at ETH tests the lifetime of the system and finds that this new computer system *cannot* guarantee a lifetime of 2 years. She writes a program to wear out the entire L-RAM device as quickly as possible. How fast is she able to wear out the device? Show all work.

$t_{wear\_out} = \frac{2^{32}}{2^3} \times 10^6 \times (40 + 4 + 20)$
$t_{wear\_out} = 2^{35} \times 10^6$ ns
$t_{wear\_out} \approx 397.68$ days

Explanation:

- Each memory cell should receive 10^6 writes.
- Since ETH-RAM is word addressable, the required number of writes is equal to $\frac{2^{32}}{2^3} \times 10^6$.
- The processor is in-order and there is no memory-level parallelism, so the total latency of each memory access is equal to $40 + 4 + 20$.

(b) [15 points] L-RAM works in the multi-level cell (MLC) mode in which each memory cell stores 2 bits. The student decides to improve the lifetime of L-RAM cells by using the single-level cell (SLC) mode. When L-RAM is used in SLC mode, the lifetime of each cell improves by a factor of 10 and the write latency decreases by 75%. What is the lifetime of the system using the SLC mode, if we repeat the experiment in part (a), with all else remaining the same in the system? Show your work.

$t_{wear\_out} = \frac{2^{31}}{2^3} \times 10^7 \times (10 + 4 + 20) \times 10^{-9}$
$t_{wear\_out} = 91268055.04\text{s} \approx 1056.34$ days

Explanation:

- Each memory cell should receive $10 \times 10^6 = 10^7$ writes.
- The memory capacity is reduced by 50% since we are using SLC: $Capacity = 2^{32}/2 = 2^{31}$
- The required number of writes is equal to $\frac{2^{31}}{2^3} \times 10^7$.
- The SLC write latency is $0.25 \times t_{write\_MLC}$: $t_{write\_SLC} = 0.25 \times 40 = 10\,\text{ns}$

(c) [10 points] Provide a mechanism that would increase the guaranteed lifetime of the computer system without changing the physical circuitry of L-RAM. From the baseline computer system in part (a), describe the changes required to guarantee a computer system lifetime of 2 years, with your mechanism. Be concrete and precise.

Artificially increase the time to either (1) send a memory request from the memory controller to L-RAM or (2) send a request from the processor to the memory controller by 54 ns. $730 * 3600 * 24 < \frac{2^{32}}{2^3} \times 10^6 \times (40 + 4 + 20 + x)$
$x > 53.4$ns