Assignment 7

Due: 1:25pm, Wed Nov 16th, 2022

Note: Make reasonable assumptions where necessary and clearly state them. Feel free to discuss problems with classmates, but the only written material that you may consult while writing your solutions are the textbook and lecture slides/videos. Solutions should be uploaded as a single pdf file on Canvas. **Show your solution steps** so you receive partial credit for incorrect answers and we know you have understood the material. Don't just show us the final answer.

This homework has an automatic penalty-free **1.5 day extension** to accommodate any covid/family-related disruptions. In other words, try to finish your homework by Wednesday 1:25pm to keep up with the lecture content, but if necessary, you may take until **Thursday 11:59pm** .

1. **Large Caches (20 points)**

   Assume a large shared LLC that is tiled and distributed on the chip. Assume that the OS page size is 16KB. The entire LLC has a size of 32 MB, uses 64-byte blocks, and is 16-way set-associative. What is the maximum number of tiles such that the OS has full flexibility in placing a page in a tile of its choosing?

2. **Virtually Indexed Cache (20 points)**

   Assume that the OS uses a minimum page size of 16 KB. Assume that your L1 cache must be 2-way set-associative. If you're trying to correctly implement a virtually indexed physically tagged cache (with no additional support from the OS or hardware), what is the largest L1 cache that you can design?

3. **Organizing Ranks (20 points)**

   Consider a system that has two processor sockets; each socket has six DDR4 memory channels. Each channel can accommodate up to four ranks. Assume that you can today purchase a DRAM chip with capacity of 2Gb, 4Gb, or 8Gb. Assume that these chips can have a data output width of 4, 8, or 16. What is the maximum capacity that can be supported by this memory system? What is the memory bandwidth supported by the system if each memory channel operates at a frequency of 1.2 GHz?

4. **Refresh (20 points)**

   Consider a memory system that has a capacity of 1 TB, that is made up of 32 ranks, each rank having 16 banks. Assume that every refresh command triggers a parallel refresh in every bank in every rank. Assume that a row in a bank has a capacity of 8 KB. Assume it takes 40 ns on average to refresh each row in a bank. Assume that every row must be refreshed within 64 ms and a refresh command is issued every 7.8 us (8,192 refresh commands are issued within a 64 ms window). How many rows are refreshed in each bank on every refresh command? For what fraction of time is the memory system unavailable performing refresh?

5. **Row Buffers (20 points)**

   For the following memory access pattern, estimate when each memory access completes for two different scheduling mechanisms: open-page policy and close-page policy. You are allowed to re-order requests already waiting in the memory controller. The access pattern only specifies the row being touched. All accesses are to the same bank. Assume that bus latencies are zero. Assume that the bank is already precharged at time 0. Assume that precharge takes 20 ns, loading a row buffer (Activate) takes 20 ns, and

cache line transfer to output pins (Column-Rd) also takes 20 ns (in other words, a row buffer hit takes 20 ns, an empty row access takes 40 ns, and a row buffer conflict takes 60 ns).

Row being accessed	Arrival time at memory controller	Open-Page	Close-Page
X	10 ns		
X	75 ns		
Y	100 ns		
X	190 ns		
X	280 ns		
Y	290 ns		