Assignment-6: Main Memory

Arnab Das, u1014840

April 18, 2018

1 Memory Scheduling Mechanisms

Row-Access	Arrival Time(ns)	open-page(ns)	closed-page(ns)
X	10	50	50
X	70	90	110
Y	90	150	170
X	100	210	230
Y	180	270	290

Table 1: Scheduling

Since the bank is already precharged, hence in both cases the first access requires loading a row buffer and cache line transfer, hence 40 second of delay. For the second access, the row is the same, hence the open page policy requires only 20ns for the cache line transfer to the output pins, while for the close-page policy it requires the time to precharge, load row buffer and cache line transfer adding a delay of 60ns. Note that for the second access for close-page, the time of arrival to the memory controller is 70ns, while it already finished the first access at 50ns. So it can utilize the 20ns gap for the precharge, and an additional 40ns for the row buffer loading and cache line transfer. For all the remaining cases, both policied require the entire cycle of precharge, row buffer loading and cache line transfer, since different rows get accessed every time. Also, reordering of request is not possible since at no point there is two existing request at the memory controller for the same array.

2 Memory System Design

Advantages of the new memory system,

- Smaller data bus increases the frequency of memory access. With a large data bus, lot of data can be read at once, but it takes larger time to collect the single burst, hence frequency of memory operations is low, whereas for larger data buses the data in a single burst is available much faster, hence frequency of memory operations increse.
- Reducing the data bus size and the accessible memory reduces the the DIMM area and the area occupied by the bus wires. It reduces

wiring cost of the communication link . Additionally, reducing wires reduces the inter-wire capacitance values , which causes the associated dynamic power to reduce.

Disadvantages of the new memory system,

- With smaller data bus, reduces the data transfer rate hence limitting the bandwidth of the operation, since smaller chunks of data are accessible on each memory access.
- Since it is connected to only four x8 memory chips, it reduces the size of the addressable memory, although since we have not changed the number of address bits, the addressable memory space remains same. One way is to multiplex the small bus size to connect to more memory chips thus utilizing the full address space while increasing the frequency of memory accesses.

3 *Virtually Indexed Physically Tagged*

If the OS uses a minimum page size of 4KB, since memory transfers in and out of the cache happens in block sizes, considering the OS doesn't intervenes, the block size can be equal to the page size. For a 8-way set-associative cache, then the size of each cache set would be 32KB. To index 32KB, it requires 15 bits of offset. If the system has N address bits, then the number of cache sets to be indexed into can be 2^{N-15}, and the total cache size can be 2^N bytes.

4 *Dram Control*

Correct row already placed in buffer If the correct row is already in buffer, then it requires only the time of the column access stobe(t_{CL}). Hence, the latency would be

$$\text{Latency } = t_{CL}$$

Another row is already placed in the row buffer In this case there is a row conflict, hence, t_{RP} is required for precharging, then t_{RCD} for the row to column delay and finally the t_{CAS}or (t_{CL}) to transfer the data onto the cache lines. Hence the latency would be

$$\text{Latency } = t_{RP} + t_{RCD} + t_{CL}$$

Since the row aaddress strobe needs to be valid for the duration of loading the data to the row-buffer and validating the column address to transfer data to the cache lines, hence $t_{RAS} = t_{RCD} + t_{CL}$. Thus, we can also write the above latency as

$$\text{Latency } = t_{RP} + t_{RAS}$$

5 *DRAM Control Tasks - Request Scheduling*

addr	rank	bank	row	column		PRE	ACT	RD
x00040108	0	2	0	108				
x01040101	0	2	10	101	→			
x00161804	0	3	1	1804				
x01040104	0	2	10	104				

Figure 1: Commands sent by scheduler

Figure 1 shows the commands to be issued by an in-order command scheduler. In total 8 commands are issued.

- Address '00040108' does not require the precharge cycle since all banks are precharged. It requires the row actiavte to load the contents into the row buffer and 'RD' cycle .
- Address '01040101' belong to the same bank as the first one but to a different row, hence it requires precharge, then row activate and then 'RD' cycle.
- Address '00161804' belong to a different bank. Hence, it is already precharged, so it requires the row actiavte and 'RD' cycle.
- Address '01040104' belongs to the same bank and row as address '01040101', hence it is a row hit, so only 'RD' cycle is necessary.