

Name:

Note: Please post your homework to ICS232 D2L on or before the due date.

Chapter 6 – Memory

Essential Terms and Concepts

4. Explain the concept of memory hierarchy. Why did the authors choose to represent it as a pyramid?

Memory hierarchy: arranging the different kinds of storage devices in a computer based on their cost, access speed, size, and roles to play in the application processing. Representing it as a pyramid allows learners to visualize the top to bottom rankings of the storage devices easier

6. What are the three forms of locality?

Temporal, spatial, and network

13. Explain how set-associative cache combines the ideas of direct and fully associative cache.

Replacement policies are established. Instead of placing memory blocks in specific cache locations based on memory address, could allow block to go anywhere in the cache. Uses multiple frames for each cache line

18. What, exactly, is effective access time (EAT)?

The performance of hierarchical memory measurement. Weighted average that takes into account the hit ratio and relative access times of the memory levels. Measurement is given by EAT = $H \times Access(c) + (1-H) \times Access(MM)$



33. What is a page fault? When logical address requires a page to be brought in from disk

Exercises

- 2. Suppose a computer using direct mapped cache has 2^{32} bytes of byte-addressable main memory, and a cache of 1024 blocks, where each cache block contains 32 bytes. $1024 = 2^{10}$ $32 = 2^{5}$
- a) How many blocks of main memory are there? 2^32/2^5=2^27 134,217,728
- b) What is the format of a memory address as seen by the cache, i.e., what are the sizes of the tag, block, and offset fields? 10 bits for block offset, 5 bits for byte offset, 17 bit for tag bits,
- c) To which cache block will the memory address 0x000063FA map? 153 blocks
- 5. Suppose a computer using fully associative cache has 2²⁴ bytes of byte-addressable main memory and a cache of 128 blocks, where each cache block contains 64 bytes.
- a) How many blocks of main memory are there? 2^24/64=2^18 256k blocks
- b) What is the format of a memory address as seen by the cache, i.e., what are the sizes of the tag and offset fields? $Log 2^2 4 = 24$ bits log 264 = 6 bits log
- c) To which cache block will the memory address 0x01D872 map? First cache block or block zero
- 16. Assume a direct-mapped cache that holds a total of 4096 bytes, where each block is 16 bytes. Assuming an address is 32 bits and that cache is initially empty, complete the table below. (You should use hexadecimal numbers for all answers.) Which, if any of the addresses will cause a collision (forcing the block that was just brought in to be overwritten) if they are accessed one right after the other?



Address	Tag	Cache location block	Offset within block
0x0FF0FABA	OFFOF	AB	A
0x0000011	0	1	1
0×0FFFFFE	OFFFF	FF	E
0x23456719	23456	71	9
0xCAFEBABE	CAFEB	AB	E

20. Suppose you have a byte-addressable virtual address memory system with 8 virtual pages of 64 bytes each, and 4-page frames. Assuming the following page table, answer the questions below:

Page #	Frame #	Valid Bit
0	1	1
1	3	0
2	-	0
3	0	1
4	2	1
5	-	0
6	-	0
7	-	0

Note: Page 1 should have the valid bit = 1.

- a) How many bits are in a virtual address? 2^3=8 2^6=4bytes 3+6=9bits
- b) How many bits are in a physical address? 2^2=4 2^6=64bytes 2+6=8



- c) What physical address corresponds to the following virtual addresses (if the address causes a page fault, simply indicate this is the case)?
 - i) 0x00 01 00 0000
 - ii) 0x44 11 00 0100
 - iii) 0xC2 00 00 0010
 - iv) 0x80 page fault occurs
- 24. Does a TLB miss always indicate that a page is missing from memory? Explain. Yes.
- 27. Consider a system that has multiple processors where each processor has its own cache, but main memory is shared among all processors.
- a) Which cache write policy would you use? Cache write policy
- b) The Cache Coherency Problem. With regard to the system just described, what problems are caused if a processor has a copy of memory block A in its cache and a second processor, also having a copy of A in its cache, then updates main memory block A? Can you think of a way (perhaps more than one) of preventing this situation, or lessening its effects?broadcasting to all caches available, where individual caches monitor address lines for access to memory locations whenever there is an update
- 29. Name two ways that, as a programmer, you can improve cache performance. Mapping technique, maintain locality of references

Prepare for next class by reading Chapter 7 – Input/Output Systems

Continue working on Project 2

Continue working on Your Group Project