

Computer Organization and Structure

Homework #4
Due: 2010/1/5

1. Caches are important to providing a high performance memory hierarchy to processors. Below is a list of 32-bit memory address references, given as word addresses.

a.	1,134,212,1,135,213,162,161,2,44,41,221
b.	6,214,175,214,6,84,65,174,64,105,85,215

- a. For each of these references, identify the binary address, the tag, and the index given a direct-mapped cache with 16 one-word blocks. Also list if each reference is a hit or a miss, assuming the cache is initially empty.
 - b. For each of these references, identify the binary address, the tag, and the index given a direct-mapped cache with two-word blocks and a total size of eight blocks. Also list if each reference is a hit or a miss, assuming the cache is initially empty.
 - c. You are asked to optimize a cache design for the given references. There are three direct-mapped cache designs possible, all with a total of eight words of data: C1 has one-word blocks, C2 has two-word blocks, and C3 has four-word blocks. In terms of miss rate, which cache design is the best? If the miss stall time is 25 cycles, and C1 has an access time of 2 cycles, C2 takes 3 cycles, and C3 takes 5 cycles, which is the best cache design?
2. In general, cache access time is proportional to capacity. Assume that main memory accesses take 70ns and that memory accesses are 36% of all instructions. The following table shows data for L1 caches attached to each of two processors P1 and P2.

		L1 size	L1 miss rate	L1 hit time
a.	P1	1KB	11.4%	0.62ns
	P2	2KB	8.0%	0.66ns
b.	P1	8KB	4.3%	0.96ns
	P2	16KB	3.4%	1.08ns

- a. Assuming that the L1 hit time determines the cycle times for P1 and P2, what are their respective clock rates?
- b. What is the AMAT for each of P1 and P2?
- c. Assuming a base CPI of 1.0, what is the total CPI for each of P1 and P2? Which processor is faster?

For the next three sub-problems, we will consider the addition of an L2 cache to P1 to presumably make up for its limited L1 cache capacity. Use the L1 cache capacities and hit times from the previous table when solving these sub-problems. The L2 miss rate indicate is its local miss rate.

	L2 size	L2 miss rate	L2 hit time
a.	512KB	98%	3.22ns

b.	4MB	73%	11.48ns
----	-----	-----	---------

- d. What is the AMAT for P1 with the addition of an L2 cache? Is the AMAT better or worse with the L2 cache?
- e. Assuming a base CPI of 1.0, what is the total CPI for P1 with the addition of an L2 cache?
- f. Which processor is faster, now that P1 has an L2 cache? If P1 is faster, what miss rate would P2 need in its L1 cache to match P1's performance? If P2 is faster, what miss rate would P1 need in its L1 cache to match P2's performance?
3. The following table is a stream of virtual addresses as seen on a system. Assume 4KB pages, a four-entry fully associative TLB, and true LRU replacement. If pages must be brought in from disk, increment the next largest page number.

a.	4095,31272,15789,15000,7193,4096,8912
b.	9452,30964,19136,46502,38110,16653,48480

TLB

Valid	Tag	Physical Page Number
1	11	12
1	7	4
1	3	6
0	4	9

Page table

Valid	Physical page or in disk
1	5
0	Disk
0	Disk
1	6
1	9
1	11
0	Disk
1	4
0	Disk
0	Disk
1	3
1	12

- a. Given the address stream in the table, and the shown initial state of the TLB and page table, show the final state of the system. Also list for each reference if it is a hit in the TLB, a hit in the page table, or a page fault.
- b. Repeat the above sub-problem, but this time use 16KB pages instead of 4KB pages. What would be some of the advantages of having a larger page size? What are some of the disadvantages?

- c. Show the final contents of the TLB if it is two-way set-associative. Also show the contents of the TLB if it is direct-mapped? Discuss the importance of having a TLB to high performance. How would virtual memory accesses be handled if there were no TLB?

There are several parameters that impact the overall size of the page table. Listed below are several key page table parameters.

	Virtual address size	Page size	Page table entry size
a.	32 bits	4KB	4 bytes
b.	64 bits	16KB	8 bytes

- d. Given the parameters in the table above, calculate the total page table size for a system running five applications that utilize half of the memory available.
- e. Given the parameters in the table above, calculate the total page table size for a system running five applications that utilize half of the memory available, given a two-level page table approach with 256 entries. Assume each entry of the main page table is 6 bytes. Calculate the minimum and maximum amount of memory required.
- f. A cache designer wants to increase the size of a 4KB virtually indexed, physically tagged cache. Given the page size listed in the table above, is it possible to make a 16KB direct-mapped cache, assuming two words per block? How would the designer increase the data size of the cache?