CS 465 – Fall 2018 – Homework 4

Prof. Daniel A. Menasce
Department of Computer Science
George Mason University

Team Allowed: maximum of two per team. The members of a team must be from the **same** CS 465 section.

State clearly the team member names and GMU IDs as comments in each page of the submitted report.

IMPORTANT:

(1) Start to work on the homework early.
(2) If you plan to do the homework with a classmate, make a determination early if your teammate is responsive.
(3) The explanation for your answers is required. **You will not get credit without explanation.**
(4) Use a word processor to type your answers to part II and generate a pdf file from it.
(5) Do not share your solutions or partial solutions on Piazza or any other platforms (electronic or otherwise) before the deadline. This assignment is individual to a group. If you want to ask questions about your solution, please ask the TAs or me.
(6) Plagiarism is not permitted in any form. We enforce the university policy and honor code

Late submissions are not accepted.

Use a Word processor to type your answers and generate a PDF file to submit your homework..

How to submit: The submission will be made via a blackboard link available to you. You must submit a PDF file answering all questions. For team projects, only one member of the team should submit the PDF file and the other should submit a one-page PDF file stating the names of both members of the team. If you fail to submit the one-page PDF file you will not get any credit for the homework.

The PDF file name should follow the following convention: "cs465_hw4_LastName1_LastName2"(for 2-member teams), or "cs465_hw4_LastName1" (for 1-member teams).

Question 1 [14 points]: Consider a byte-addressable memory with a single level of cache that can store 1024 4-word blocks. All questions below are relative to this memory/cache system.

(a) Assume a direct mapped cache organization. How many bits are used for the cache index, block/byte offset, and tag?

(b) Still assuming a direct mapped cache organization, how many bits are needed to store all the tags in the cache?

(c) Assume a 2-way set associative mapped cache organization. How many bits are used for the cache index, block/byte offset, and tag?

(d) Still assuming a 2-way set associative mapped cache organization, how many bits are needed to store all the tags in the cache?

(e) Assume a 4-way set associative mapped cache organization. How many bits are used for the cache index, block/byte offset, and tag?

(f) Still assuming a 4-way set associative mapped cache organization, how many bits are needed to store all the tags in the cache?

(g) Assume a fully associative cache organization. How many bits are used for the cache index, block/byte offset, and tag?

(h) Still assuming a fully associative cache organization, how many bits are needed to store all the tags in the cache?

Question 2 [12 points]: Consider a processor with a single level of cache. The base CPU CPI is 1 cycle, the cache miss rate per instruction is 4% and the access time to main memory is 100 cycles. What is the average memory access time in cycles?

Question 3 [12 points]: Consider that you add a second level of cache to the processor of the previous question. The access time to level 2 cache is 5 cycles and the miss rate per instruction at the level 2 cache is 2%. What is the average memory access time in cycles?

Question 4 [12 points]: Consider that you add a third level of cache to the processor of the previous question. The access time to level 3 cache is 15 cycles and the miss rate per instruction at the level 3 cache is 0.5%.

(a) What is the average memory access time in cycles?

(b) Was the addition of the third level of cache significantly beneficial?

(c) Is it possible to bring down the average memory access time to 1.1 cycles by decreasing the third level cache access time?

Question 5 [25 pts]: Consider a virtual memory system that has a 4-entry fully associative TLB, a virtual address space that consists of 12 virtual pages (numbered 0 to 11). Physical memory can hold 6 pages (page frames are numbered 0 to 5). The current state of the TLB

and of the Page Table are shown below.

TLB

Valid Bit	Tag (VPN)	Physical Frame Number
1	0	5
1	5	2
1	2	1
0	4	3

Page Table

Virtual Page Number	Valid Bit	Physical Frame Number or on Disk
0	1	5
1	0	Disk
2	1	1
3	0	Disk
4	1	0
5	1	2
6	0	Disk
7	1	3
8	0	Disk
9	0	Disk
10	1	4
11	0	Disk

Main Memory

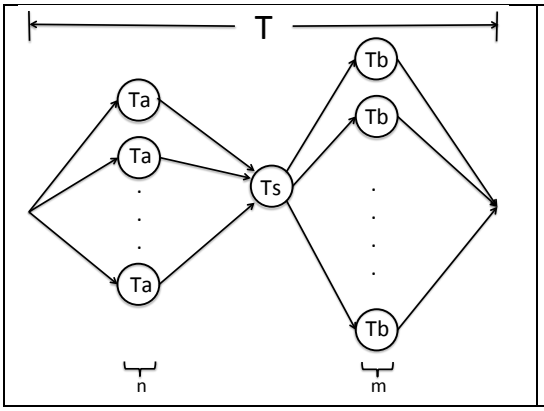
Page Frame Number	Virtual Page Number	Valid
0	4	1
1	2	1
2	5	1
3	7	1
4	10	1
5	0	1

Consider the following virtual page reference string (5, 4, 9, 7) and fill out the table below and indicate by using slashes how values in any of the tables above change (e.g., 4 /5 / 0 indicates that an initial value of 4 changed to 5 and then to 0, Disk / 2 indicates that Disk changed to 2).

Use the following page replacement policy: use perfect LRU within the page reference string above (i.e., 5, 4, 9, 7) and for pages that are not part of that reference string consider the highest virtual page number as the least recently used.

Virtual Page Number	TLB Miss or Hit?	Page Fault (Yes/No?)	Physical Page Number
5			
4			
9			
7			

Question 6 [25 pts]: This question is about running parallel applications using cloud resources. Consider the figure below that shows a parallel application that consists of three processing phases: (1) *n* identical tasks execute in parallel, each requiring Ta units of computation, (2) a sequential task that executes after all tasks of phase 1 complete and requires Ts units of computation, (3) *m* identical tasks execute in parallel requiring Tb units of computation each; these tasks can only start after the sequential task completes. Consider also, that the tasks run on virtual machines leased from a cloud provider (e.g., Amazon's Elastic Compute Cloud). The table below shows four VM types available from the cloud provider, the price/sec for each one of them, and their relative speed. For example, VM type 1 is the slowest and VM type 4 is 8 times faster than VM type 1. The times Ta, Ts, and Tb are relative to VMs of type 1. Let S(i) be the relative speed of VM of type i and let P(i) be the price/sec for VMs of type i. For example, S (3) = 4 and P (3) = $0.000110556/sec.



VM type	price/sec	Relative speed (computations/sec)
1	0.000027778	1
2	0.000055278	2
3	0.000110556	4
4	0.000221111	8

Answer the following questions:

(1) Give an expression for the total time T to execute the parallel application described above assuming that VMs of type *s* are used for phase 1, a VM of type *t* is used for the sequential task, and VMs of type *u* are used for phase 3. Your expression should be a function of Ta, Ts, Tb, S(s), S(t), and S(u).

(2) Give an expression for the total cost C to execute the parallel application described above assuming that VMs of type *s* are used for phase 1, a VM of type *t* is used for the sequential task, and VMs of type *u* are used for phase 3. Your expression should be a function of Ta, Ts, Tb, S(s), S(t), S(u), m, and n.

(3) Assume that n = 200, m = 150, Ta = 25,000 sec, Ts = 100,000 sec, and Tb = 50,000 sec. Use the numbers in the table above to compute the execution time and corresponding cost of the application using the fastest VMs.

(4) Using the same numbers as above, compute the execution time and corresponding cost of the application using the slowest VMs.

(5) Compare your answers to items (3) and (4) above and draw conclusions.