Assignment Question 1 (4 + 3 Marks):

a) A thread switch should be much more efficient than a process switch, which typically requires hundreds to thousands of processor cycles while a thread switch can be instantaneous. Why?

Sample Solution: Threads shares memory but process does not. Process need to save it's context (registers, OS state), load new process's context into memory (memory flushed and loaded). But for thread it is just change in the value of the program counter (PC) no memory flush and load since they share the memory.

b) Suppose you want to perform two sums: one is a sum of 20 scalar variables (assume can not parallelize), and one is a matrix sum of a pair of two-dimensional arrays, with dimensions 20 by 20. What speed-up do you get with 100 processors? Show the impact on speed-up if one processor's load is twice than all the rest.

Sample Solution: Similar to example in Page 506, Page 508

Assignment Question 2 (2 + 2 Marks):

The following code is written in C, where elements within the same row are stored contiguously.

```
int size = 10;
int outer, inner;
for(outer =1; outer< size; outer++)
        for(inner=1;inner<size; inner++)
                rMat[outer][inner]=lMat[outer][1] + rMat[inner][outer]
```

a) Which variables exhibit temporal locality?
b) Which variables exhibit spatial locality?

Sample Solution:

a) lMat[outer][1]
b) rMat[outer][inner]

Assignment Question 3 (4 + 4 + 4 Marks):

Below is a list of 32-bit memory address references, given as word addresses. 3, 180, 43, 2, 191, 88, 190, 14, 181, 44, 186, 253

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.

Table 1

Word address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	3	M
180	1011 0100	11	4	M
43	0010 1011	2	11	M
2	0000 0010	0	2	M
191	1011 1111	11	15	M
88	0101 1000	5	8	M
190	1011 1110	11	14	M
14	0000 1110	0	14	M
181	1011 0101	11	5	M
44	0010 1100	2	12	M
186	1011 1010	11	10	M
253	1111 1101	15	13	M

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 8 blocks. Also list if each reference is a hit or a miss, assuming the cache is initially empty.

Table 2

Word address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	1	M
180	1011 0100	11	2	M
43	0010 1011	2	5	M

2	0000 0010	0	1	H
191	1011 1111	11	7	M
88	0101 1000	5	4	M
190	1011 1110	11	7	H
14	0000 1110	0	7	M
181	1011 0101	11	2	H
44	0010 1100	2	6	M
186	1011 1010	11	5	M
253	1111 1101	15	6	M

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 8 words of data: C1 has 1-word blocks, C2 has 2-word blocks, and C3 has 4-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 (**show the details calculation**)?

Cache 1

Word address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	3	M
180	1011 0100	22	4	M
43	0010 1011	5	3	M
2	0000 0010	0	2	M
191	1011 1111	23	7	M
88	0101 1000	11	0	M
190	1011 1110	23	6	M
14	0000 1110	1	6	M
181	1011 0101	22	5	M
44	0010 1100	5	4	M
186	1011 1010	23	2	M
253	1111 1101	31	5	M

Cache2

Word address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	1	M
180	1011 0100	22	2	M

43	0010 1011	5	1	M
2	0000 0010	0	1	M
191	1011 1111	23	3	M
88	0101 1000	11	0	M
190	1011 1110	23	3	H
14	0000 1110	1	3	M
181	1011 0101	22	2	H
44	0010 1100	5	2	M
186	1011 1010	23	1	M
253	1111 1101	31	2	M

Cache 3

Word address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	0	M
180	1011 0100	22	1	M
43	0010 1011	5	0	M
2	0000 0010	0	0	M
191	1011 1111	23	1	M
88	0101 1000	11	0	M
190	1011 1110	23	1	H
14	0000 1110	1	1	M
181	1011 0101	22	1	M
44	0010 1100	5	1	M
186	1011 1010	23	0	M
253	1111 1101	31	1	M

Cache 1 miss rate = 100%
Cache 1 total cycles = 12 * 25 + 12 * 2 = 324

Cache 2 miss rate = 10/12 = 83%
Cache 2 total cycles = 10 * 25 + 12 * 3 = 286

Cache 3 miss rate = 11/12 = 92%
Cache 3 total cycles = 11 * 25 + 12 * 5 = 335

Cache 2 provides the best performance.

Assignment Question 4 (4 + 4 + 4 Marks):

Below is a list of 32-bit memory address references, given as word addresses. 3, 180, 43, 2, 191, 88, 190, 14, 181, 44, 186, 253

a) Show the final cache contents for a three-way set associative cache with two-word blocks and a total size of 24 words. Use LRU replacement. For each reference identify the index bits, the tag bits, the block off set bits, and if it is a hit or a miss.

Sample Solution: The cache would have 24/3 = 8 blocks per way and thus an index field of 3 bits.

Word Address	Binary Address	Tag	Index	Hit/Miss	Way 0	Way 1	Way 2
3	0000 0011	0	1	M	T(1)=0		
180	1011 0100	11	2	M	T(1)=0 T(2)=11		
43	0010 1011	2	5	M	T(1)=0 T(2)=11 T(5)=2		
2	0000 0010	0	1	M	T(1)=0 T(2)=11 T(5)=2	T(1)=0	
191	1011 1111	11	7	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11	T(1)=0	
88	0101 1000	5	4	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5	T(1)=0	
190	1011 1110	11	7	H	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5	T(1)=0	
14	0000 1110	0	7	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5	T(1)=0 T(7)=0	
181	1011 0101	11	2	H	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5	T(1)=0 T(7)=0	

44	0010 1100	2	6	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5 T(6)=2	T(1)=0 T(7)=0	
186	1011 1010	11	5	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5 T(6)=2	T(1)=0 T(7)=0 T(5)=11	
253	1111 1101	15	6	M	T(1)=0 T(2)=11 T(5)=2 T(7)=11 T(4)=5 T(6)=2	T(1)=0 T(7)=0 T(5)=11 T(6)=15	

b) Show the final cache contents for a fully associative cache with one-word blocks and a total size of 8 words. Use LRU replacement. For each reference identify the index bits, the tag bits, and if it is a hit or a miss.

Sample Solution: Fully associative and one-word block. The word address is equivalent to the tag. The only possible way for there to be a hit is a repeated reference to the same word, which is not the case here. So, no hit.

Tag	Hit/Miss	Contents
3	M	3
180	M	3, 180
43	M	3, 180, 43
2	M	3, 180, 43, 2
191	M	3, 180, 43, 2, 191
88	M	3, 180, 43, 2, 191, 88
190	M	3, 180, 43, 2, 191, 88, 190
14	M	3, 180, 43, 2, 191, 88, 190, 14
181	M	181, 180, 43, 2, 191, 88, 190, 14
44	M	181, 44, 43, 2, 191, 88, 190, 14
186	M	181, 44, 186, 2, 191, 88, 190, 14
253	M	181, 44, 186, 253, 191, 88, 190, 14

c) What is the miss rate for a fully associative cache with two-word blocks and a total size of 8 words, using LRU replacement? What is the miss rate using MRU (most recently used) replacement? Finally what is the best possible miss rate for this cache, given any replacement policy?

Address	Tag	Hit/ Miss	Contents
3	1	M	1
180	90	M	1, 90
43	21	M	1, 90, 21
2	1	H	1, 90, 21
191	95	M	1, 90, 21, 95
88	44	M	1, 90, 21, 95, 44
190	95	H	1, 90, 21, 95, 44
14	7	M	1, 90, 21, 95, 44, 7
181	90	H	1, 90, 21, 95, 44, 7
44	22	M	1, 90, 21, 95, 44, 7, 22
186	143	M	1, 90, 21, 95, 44, 7, 22, 143
253	126	M	1, 90, 126, 95, 44, 7, 22, 143

The final reference, 253 (126) replaces tag 21, since tag 1 and 90 has been recently used.

Miss rate = 9/12 = 75%

Even if we use the MRU tag 143 would be replaced by 126 but there would be no change in the miss rate.