1. (12%) Consider the following processes:

Process	Arrival Time	Burst Time
P1	0	6
P2	2	4
P3	3	2
P4	5	8

a) Suppose we are using a non-preemptive Shortest-Job-First scheduler. (i)Write the execution order of these processes. (ii)What is the average waiting time?

b) Now suppose we are using a preemptive Round-Robbin scheduler with a time quantum of 4 units. (i)Write the execution order of these processes. (ii)What is the average waiting time?

2. (4%) What are the strategies for dealing with deadlocks.

3. (9%) Consider three computers in the same collision domain as shown in the following figure.

Computer A
IP: 140.114.79.3
MAC: 04-10-D5-98-22-AA

Computer B
IP: 140.114.79.5
MAC: 04-10-D5-98-22-BB

Computer C
IP: 140.114.79.7
MAC: 04-10-D5-98-22-CC

Initial ARP table:

Computer A

IP	MAC
140.114.79.254	00-90-7f-0f-d2-7d

Computer B

IP	MAC
140.114.79.254	00-90-7f-0f-d2-7d

Computer C

IP	MAC
140.114.79.254	00-90-7f-0f-d2-7d

What are contents of the ARP tables of computers A, B, and C after computer A sends a packet to computer C?

4. (15%) The terms *big-endian* and *little-endian* were originally found in Jonathan Swift's book, Gulliver's Travels. Now all processors must be designated as either big-endian or little-endian. For example, DEC Alpha RISC and Intel 80x86 processors are little-endian. Motorola 680x0 microprocessors and Sun SuperSPARC are big-endian.

a) (5%) Briefly explain the differences between *big-endian* and *little-endian*.

b) (10%) Please illustrate *big-endian* and *little-endian* by considering the number 4097 stored in a **4-byte** integer.

Address	Big-Endian representation	Little-Endian representation
00		
01		
02		
03		

5. (10%) A computer whose processes have 1024 pages in their address spaces keeps its page tables in memory. The overhead required for reading a word from the page table is 500 nsec. In order to reduce the overhead, the computer has a TLB (Translation Look-aside Buffer), which holds 32 (virtual page, physical page frame) pairs, and can do a look up in 100 nsec. What hit rate is needed to reduce the mean overhead to 200 nsec?

6. (7%) Assume the instruction "ADD R0, R1, R2, LSL#2" in one instruction set architecture conducts R0=R1+R2×4 operation. Could you conduct R0=99×R1 using two ADD instructions? Write down the code if your answer is YES; otherwise, state the reasons.

7. (12%) Forwarding is a technique to eliminate the data hazards occurred among the pipelining instructions. However, not all data hazards can be handled by the forwarding. If an instruction following a LOAD instruction depends on the results of the LOAD instruction, the data hazard is occurred, and the pipeline is stalled one cycle.

   (a) (3%) Assume the percentage of LOAD instruction is 20% in a program, and half the time the instruction following a LOAD instruction needs the result of the LOAD instruction. What is the performance degradation due to the data hazard.

   (b) (6%) Pipeline scheduling or instruction scheduling techniques could be used to eliminate the data hazard mentioned in (a). What is the philosophy of pipeline scheduling? Use an example to demonstrate that pipeline scheduling can eliminate the data hazard.

   (c) (3%) What is the possible overhead of pipeline scheduling?

8. (6%) State two reasons that MIPS is not an accurate measure for comparing performance among computers.

9. (15%)

  a) (10%) Consider the following sequence of address references given as word addresses:

  　　　22, 10, 26, 30, 23, 18, 10, 14, 30, 11, 15, 19

  For a 2-way set associative cache with a block size of 8 bytes, a word size of 4 bytes, a data capacity of 64 bytes and the LRU replacement, label each reference in the sequence as a hit or a miss. Assume that the cache is initially empty.

  b) (5%) Determine the number of bits required in each entry of a TLB that has the following characteristics:

  　□ The TLB is directed-mapped

  　□ The TLB has 32 entries

  　□ The page size is 1024 bytes

  　□ Virtual byte addresses are 32 bits wide

  　□ Physical byte addresses are 31 bits wide

  Note that you only need to consider the following items for each entry:

  　□ The valid bit

  　□ The tag

  　□ The physical page number

10. (10%)

   a)   (4%) The average memory access time (AMAT) is defined as

$$\text{AMAT} = \text{time for a hit} + \text{miss rate} \times \text{miss penalty}$$

   Consider the following two machines:

      ☐Machine 1: 100 MHz, a hit time of 1 clock cycle, a miss rate of 5% and a miss penalty of 20 clock cycles

      ☐Machine 2: 100 MHz, a hit time of 1.2 clock cycles, a miss rate of 3% and a miss penalty of 25 clock cycles

   Determine which machine has smaller AMAT.

   b)   (3%) Assume that you are running a program which uses a lot of data and that 50% of the data the program needs causes page faults and must be retrieved from a disk array. If the program needs data at the rate of 600 Mbytes/second and each disk in the disk array can supply data at the rate of 30 Mbytes/second, what is the minimum number of disks required in the disk array? You do not need to worry about disk errors.

   c)   (3%) Assume that there are 10 pairs of processors and disk arrays placed all over a network. Each processor needs data at the rate of 600 Mbytes/second from its disk array across the network. If one-third of the total traffic crosses the bisection of the network, what is the bisection bandwidth needed (in Mbytes/second)?