NATIONAL EXAMINATIONS MAY 2011
98-COMP-A5 OPERATING SYSTEMS
3 Hours Duration

NOTES:

1. If doubts exist as to the interpretation of any question, the candidate is urged to submit with the answer paper a clear statement of any assumption made.

2. Provide justifications for your answers. Show all your work.

3. CLOSED BOOK. Candidates may use one of the two calculators, a Casio or Sharp approved model. No other aids.

4. The candidate has to answer **any five questions** (each question has multiple parts).

5. Total Marks = 100.

6. This exam has got **4 pages** (including this page).

1 [20 marks]

(a) Consider the two following methods for allocation of pages to programs: *fixed* and *variable*. For a fixed allocation the number of frames allocated to the program does not change whereas for a variable allocation the number of pages allocated to a program can increase or decrease during the lifetime of the program. Discuss the advantages and shortcomings (if any) for each of these approaches.

(b) Consider the following page reference string on a demand paged virtual memory system:

511,512,513,514,512,511,515,516,512,511,512,513,517,516,513,512,511,512,513,516

(i) What is the minimum number of page faults that can be achieved on the system for this reference string? Which memory management policy will be able to achieve this minimum number of page faults? Include parameter values (if any) that this policy needs to use for producing the minimum number of page faults.

(ii) What is the minimum number of page faults for the page reference string described earlier when a fixed allocation is used and 3 frames are allocated to the program?

(c) Briefly distinguish between compile time and load time address binding in the context of memory management.

2 [20 marks]

(a) Consider a **preemptive** short term scheduling strategy in which the priority of a process may change dynamically with time. (Larger priority numbers imply higher priority). At any point in time the highest priority process is run on the system. Ties are broken in favour of the process that entered the ready to run queue first. If a running process is preempted then its time of entry into the ready to run queue is the time at which the preemption was made. Assume that each process has a single CPU burst.

The priority of all processes is set to **P** when they enter the ready to run queue (upon arrival or after being preempted). When a process is waiting in the ready to run queue its priority changes at a rate **a**. That is,

Priority of a process in the ready to run queue = P + at
(where t is the time elapsed (in seconds) after the process entered the ready to run queue).

When a process is selected to run on the CPU its priority changes at a rate **b**. That is,

Priority of the process running on the CPU = Old Priority (when allocated CPU) + bt'
(where t' is the time elapsed (in seconds) after the process started running on the CPU).

The parameters P, a, and b can be set to give many different scheduling policies. Once chosen the values of these parameters become fixed and can not change.

(i) Determine P, a, and b that will produce the First Come First Served policy.
(ii) Determine P, a, and b that will produce the Last Come First Served policy. Under this policy whenever a process arrives on the system it preemptively captures the CPU. That is,

if the CPU is free the process is allocated the CPU; if the CPU is busy then the executing process is preempted and the CPU is allocated to the process that just entered the ready to run queue. Whenever a process completes and the CPU is free the process (in the ready to run queue) that was preempted most recently is allocated the CPU.

(b) For a system that controls a nuclear reactor what kind of a system should one devise - a conventional time-sharing system, a hard real time system or a soft real time system?

3. [20 marks]

(a) Distinguish between a deadlock state and an unsafe state. Using an example explain how an unsafe state may not necessarily lead to a deadlock state.

(b) Discuss the *deadlock prevention* approach that can be deployed on a system to prevent the occurrence of a deadlock. Include the limitations (if any) of the approach in your discussion.

(c) Consider a moving head hard disk which consists of a single platter (surface) with 180 tracks on it. The tracks are numbered 0 to 179. The disk is currently serving a request at track 140 and has just finished a request at track 135. The queue of pending requests in FIFO order is:

86, 147, 91, 177, 94, 150, 102, 175, 130.
[Each number indicates the track number corresponding to a request]

What is the total head movement (in number of tracks) needed to satisfy all these requests for the following disk scheduling algorithms?

(i) Shortest Seek Time First (SSTF) (ii) LOOK

4. [20 marks]

(a) Discuss how file sharing is achieved on a system. Include the problems associated with the file sharing technique in your discussion.

(b) Discuss why controlling accesses to files are important in the context of multi-user systems. Describe any two methods of maintaining access rights for files on the system and how they are used for access control.

(c) Different methods are available for the management of blocks allocated to files. Consider a file currently consisting of 150 blocks. Assume that the directory is available in main memory). The 55^{th} block in the file needs to be deleted. For each of the following cases compute the minimum number of disk operations that are required. Assume that there is no room for the file to grow in the beginning but there is room to grow in the end.

(i) contiguous allocation (ii) linked allocation (based on a singly linked list) (iii) indexed allocation (assume that only one index block is used)

Note that each disk operation corresponds to the reading of a block from the disk or the writing of a block to the disk. While computing the number of disk operations ignore the disk operations that may be required for the location and maintenance of free space. Since the directory is in main memory any operation on the directory is not counted as a disk operation.

5. [20 marks]

(a) Discuss how the three requirements associated with the solution to the critical section problem are satisfied when a *monitor* is used for protection of shared data.

(b)Briefly describe the "Readers-Writers" problem that is well known in the context of concurrent processes. Provide the algorithm for a single monitor-based solution to the "Readers-Writers" problem.

(c) Explain if using more than one monitor for its solution can increase the degree of concurrency in process execution. You do not need to provide an algorithm for this case even if your answer is yes. A short answer followed by an appropriate justification is sufficient.

6. **[20 marks]**

(a) Discuss whether increasing the system memory can improve the CPU utilization in the following scenarios. If your answer is NO for any one of these scenarios describe what change needs to be made for the improvement of CPU utilization.

(i) when the CPU utilization is 10% and the paging disk utilization is 5%
(ii) when the CPU utilization is 10% and the paging disk utilization is 95%
(iii) when both the CPU and paging disk utilizations are 60%

(b) What is thrashing? Describe with the help of an example a memory management technique that can effectively control thrashing.

(c) Distinguish between external and internal fragmentation in the context of memory management.
Discuss the impact of (i) paging and (ii) compaction on each type of fragmentation.
Include the overhead associated with each of these techniques in your discussion.

7. [20 marks]

(a) Explain how multiple CPUs and multiple disks can be used to enhance both system reliability and performance. Consider both sequential as well as parallel applications in your discussion.

(b) Briefly describe a technique used by the operating system for protection of data sent over a communication link. Discuss the overheads associated with the technique.

(c) Discuss the bit-vector and the linked list based techniques for managing free space on disk systems. Include the merits and shortcomings of each technique in your discussion

(d) Briefly describe the direct and sequential access techniques used in the storage and retrieval of data. Provide examples of an application that performs sequential access of data and of another application that performs random access of data.