UNIVERSITY OF MORATUWA

FACULTY OF ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

MSc in Computer Science
2015 Intake Semester 2 Examination

CS5202 ADVANCED OPERATING SYSTEMS

Time allowed: 2 Hours

August 2015

ADDITIONAL MATERIAL: *None*

INSTRUCTIONS TO CANDIDATES:

1. This paper consists of **4** questions in **6** pages.

2. Answer **all** questions.

3. Please note that the total marks allocated to questions are different and indicated at the beginning of each question.

4. Start answering each of the main questions on a new page.

5. The maximum attainable mark for each question is given in brackets.

6. This examination accounts for 50% of the module assessment.

7. This is a closed book examination.

    NB: It is an offence to be in possession of unauthorized material during the examination.

8. Only calculators approved by the Faculty of Engineering are permitted.

9. Assume reasonable values for any data not given in or with the examination paper. Clearly state such assumptions made on the script.

10. In case of any doubt as to the interpretation of the wording of a question, make suitable assumptions and clearly state them on the script.

11. This paper should be answered only in English.

Question 1 (25 marks)

A Select the most appropriate answer, and write the corresponding sub question number and the answer number in your answer book. [10 × 1]

(i) Which of the following is a goal of an operating system?

A) Act as a resource allocator   B) Control execution of programs

C) Provide an abstraction to hardware   D) All of these

(ii) A user program executes in a _____ , in which certain areas of memory are protected from the user's use, and in which certain instructions may not be executed.

A) batch mode   B) kernel mode

C) task mode   D) user mode

(iii) There are 3 processes P, Q, and R with execution times of 10 ms, 4 ms, and 8 ms, respectively. All the processes arrive at the same time in the order of P, Q, and R. What is the waiting time for process R, if the scheduling algorithm is shortest job first?

A) 0   B) 4

C) 8   D) 10

(iv) For the same set of processes in question (iii), what is the average turnaround time, if the scheduling algorithm is first come first serve?

A) 4   B) 38/3

C) 50/3   D) 22

(v) Which of the following statements are TRUE about Multithreading models?

(p) One-to-One model can benefit from hardware-level parallelism
(q) Two-level model allows a user thread to be bound to a kernel thread
(r) One-to-Many model allows the OS to create any no of kernel threads

A) (p) and (q) only   B) (p) and (r) only

C) (q) and (r) only   D) All three

(vi) Which of the following statements are TRUE about Semaphores?

(p) Enable mutual exclusion and process synchronization
(q) Semaphores do not require busy waiting
(r) Value of a binary semaphore change only between 0 and 1

A) (p) and (q) only   B) (p) and (r) only

C) (q) and (r) only   D) All three

(vii) One technique for overcoming internal fragmentation of memory is _____ .

A) having large pages   B) having small pages

C) relocation of pages   D) same sized pages and frames

(viii)   Which of the following is **not** a typical memory segment in a program's memory layout?

A) code

B) heap

C) pool

D) stack

(ix)   In the Dynamic Partitioning technique of memory management, the placement algorithm that chooses the block that is closest in size to the request is called

A) best-fit

B) first-fit

C) last-fit

D) next-fit

(x)   In rate monotonic scheduling

A) longer duration jobs have higher priority

B) shorter duration jobs have higher priority

C) priority does not depend on the duration of the job

D) none of the above

B   State whether the following statements are TRUE or FALSE. Give one sentence justification for your answer.   [5 × 2]

(i)   In a time sharing system, a user's program is preempted at regular intervals, but due to relatively slow human reaction time this occurrence is usually transparent to the user.

(ii)   When physical memory becomes free a process move from Blocked/Suspend to Blocked state.

(iii)   First Come First Serve performs much better for short processes than long ones.

(iv)   All segments of a program must be of the same length.

(v)   Hardware Abstraction Layer (HAL) presents a consistent API to the upper layers and maps upper-layer operations onto a specific hardware platform.

C   Write the most appropriate short answer (word/phrase) for following questions. [5 × 1]

(i)   In _____ CPU continuously wait for a task to be performed.

(ii)   _____ architectures allow more than 1 instruction to execute during a clock cycle by simultaneously dispatching multiple instructions to redundant functional units.

(iii)   A situation in which 2 or more processes are waiting indefinitely for an event that can be caused by only one of the waiting processes is a _____ .

(iv) _____ brings a page into physical memory only when it is essential.

(v) _____ occurs when sum of localities of all running programs exceed the physical memory size.

Question 2 (25 marks)

(i) Discuss whether a micro or monolithic kernel based design is more suitable for the following requirements. Justify your selection.

   a) An operating system designed to run on multiple devices such as desktops, servers, smart phones, and game consoles. [4]

   b) An operating system designed for a modern car, which includes engine control, breaking, driver assistance, navigation, and rich multimedia experience. [4]

(ii) "It is recommended to use Kernel Threads to benefit from modern multi-core processors"

Do you agree or disagree with this statement? Justify your decision. [4]

(iii) Following 2 options are suggested for a 64-bit memory system. Discuss the advantages and disadvantages of each of the following design choices.

   a) A logical address with a 20-bit page offset and a 2-level page table. [4]

   b) A logical address with a 32-bit page offset and a hashed page table. [4]

(iv) Following design is proposed for a virtual memory system of an operating system designed for smart phones and tablets.

"Proportional allocation of memory frames based on process priorities with local replacement policy for frames"

Critically evaluate the suitability of this design for the said operating system. [5]

Question 3 (25 marks)

One of your clients is into the business of selling toys and baby items. The client recently deployed a new web based, Point of Sales (POS) system across multiple stores in major cities. While the POS provides the necessary business functions, client is experiencing significant delays in processing transactions and updating inventory. These issues are noticed only during certain times of the day. According to the client following are the symptoms observed:

"Our server has 4 Intel Xeon cores running at 3.2 GHz. Server has 16 GB of RAM and 4 TB of disk space. POS database is also on the same node. Server runs Windows Server 2012 R2. Web and database servers are IIS and Microsoft SQL server, respectively. Each branch has 1 Mbps leased line and head office has a 10 Mbps leased line. To provide secure communication POS web server use HTTPS connections.

Sometimes it takes several seconds to search for an item once the barcode is scanned. It takes tens of seconds for activities like accepting payments, printing receipts, updating inventory, and updating customer's loyalty card points/balance. We have realized the CPU utilization of the server is mostly below 55%. It occasionally jumps to 100% when we run batch updates, e.g., generating end-of-day summary. However, those updates are shot lived and carried out mostly after closing the stores at 9:00 pm. Average number of processes or threads in the ready queue is 1.82 per CPU core. However, average number of process or threads in the waiting queue is 23.3 per CPU core. Average memory utilization is 68% while peak memory utilization is 100%. Memory peaks tends to occur at random times of the day; however, they do not last more than 2.6 seconds in average. Disk utilization is 25%. However, we see peak read and write IOPS (I/O Operations per Second) as high as 243,400 and 56,700, respectively. Most of the disk reads and writes are short lived. However, long reads and writes that last tens of seconds are not uncommon. We also observed several locks on certain database entries which last few seconds."

Client want your feedback to diagnose the issue(s) of the POS system and want your input on how to resolve the issue(s). Answer following questions based on this scenario.

(i) To improve performance of POS system, should we give high priority to CPU bound processes/treads or I/O bound processes/threads? Justify. [5]

(ii) Using the given case as an example, briefly describe what causes a process to move from running to waiting queue (i.e., running to blocked state). [3]

(iii) Using the given case as an example, discuss how the following pairs of scheduling criteria/objectives conflict in certain settings.

a) CPU utilization and response time [3]

b) I/O device utilization and response time [3]

(iv) One of your colleagues suggests to run multiple instances of the web server to improve the throughput and reduce latency as the CPU utilization is not very high.

Do you agree or disagree to this suggestion? Briefly explain. [3]

(v) Another suggestion is to increase the number of CPU cores to 8, RAM to 32 GB, and also have a fast disk system with RAID (Redundant Array of Independent Disks). Will this make the workload change its characteristics from CPU bound to I/O bound or vice versa? Discuss. [4]

(vi) Another suggestion is to port the POS application to an Exokernel based operating system as it gives untrusted software efficient control over hardware and software resources by separating management from protection. Also, it is demonstrated that a webserver on top of an Exokernel-based operating system can gain 8× performance.

Do you agree or disagree to this suggestion? Briefly explain. [4]

Question 4 (25 marks)

Answer **any 5** questions. These questions are based on papers discussed in the class.   [5 × 5]

(i) Following is a diagram of a priority-based multilevel feedback model. Describe what happens, if the execution time t of a newly arriving process with *priority 0* is much larger than the quantum size q (i.e., $t >> q$).
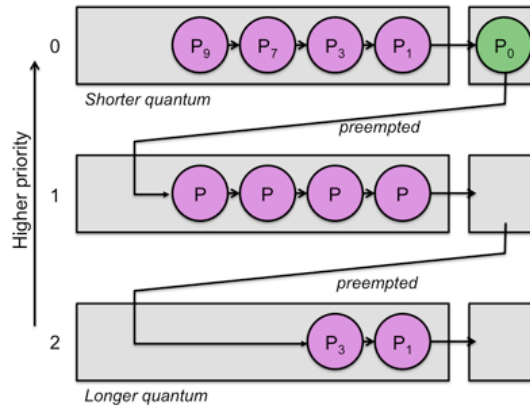
(ii) "Frequent allocation and freeing of memory creates holes within the memory allocated to a process. This leads to inefficient memory utilization."

Using a suitable diagram briefly describe how frequent allocation and freeing of memory creates holes within the memory and how log-structured memory overcome this issue.

(iii) "Google File System (GFS) is not designed to be a general-purpose file system and is optimized for large files (rather than lots of small files), streaming reads (more common than writes), and atomic appends (which may be done concurrently by multiple clients)"

Using a suitable diagram discuss how the GFS architecture is designed to achieve the above objectives.

(iv) "Each scheduling decision in Apollo (a highly scalable and coordinated scheduling framework for cloud computing) considers future resource availability and optimizes various performance and system factors together in a single unified model."

Briefly describe how Apollo determines future resource availability and how such knowledge/prediction leads to better scheduling decisions.

(v) "Windows is known to have better multimedia performance than Linux. This gap is expected to widen with DirectX version 12 included in Windows 10. DirectX is an API for multimedia related operations like game programming, and video. DirectX 12 provides a vastly better threading model, increased performance, and up to 50% power savings."

By considering the architectural aspects of Windows family of operating systems discuss how such performance can be gained.

(vi) Mobile processors with heterogeneous CPU cores create new problems of cache coherence.

Using a suitable example explain what is cache coherence. Also, briefly outline a solution to address this problem where a mobile app can seamlessly switch between different CPU cores to gain power-performance advantages.

-------------------------- END OF THE PAPER --------------------------