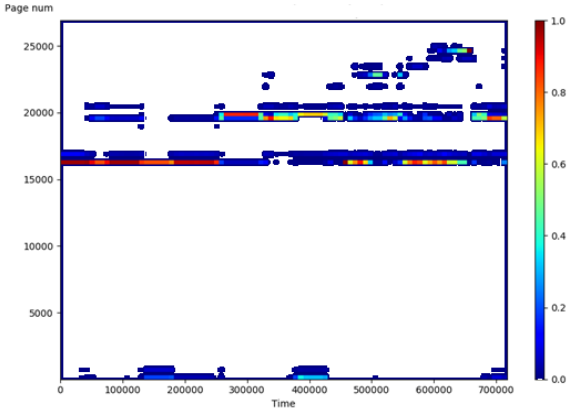
Introduction to Operating Systems, 2022 Fall
Question Sheet of Final

1. [C7, 9 pts] Consider a set of processes that have been deadlocked. Can each of the following possibly break the deadlock? Why or why not?
    a. Increasing the CPU clock rate
    b. Killing some processes
    c. Changing the CPU scheduling algorithm
2. [C7, 6 pts] A mutex lock is very similar to a semaphore with an initial value of 1. However, while a semaphore can be signaled by any process, a mutex lock can only be unlocked by its owner. Explain why there is such a difference.
3. [C8, 6 pts] The problem of dynamic memory allocation is inherently intractable, i.e., the optimal solution does not exist. Among many excellent heuristics, Best-Fit (BF) and First-Fit (FF) are two popular ones. Demonstrate an allocation-deallocation sequence for each of the following cases:
    a. BF succeeds the sequence but FF fails
    b. FF succeeds the sequence but BF fails
4. [C8, 8 pts] A (forward) page table entry contains a frame number and a set of status bits, which are set/cleared by the MMU. Explain the meaning of the following page status bits and discuss an application of each of the following bits:
    a. Dirty bit
    b. Reference bit
    c. Read-only bit
    d. Valid (present) bit
5. [C8, 3 pts] Consider a microprocessor whose virtual addresses and physical addresses are both 32 bits. The microprocessor uses two-level paging, for which the 1st-level page directory has 1024 entries; it also supports page sizes of 4KB and 4MB. Answer the following:
    a. What is the benefit of using large pages?
    b. What is the benefit of using small pages?
    c. Demonstrate how the two-level paging supports both 4KB and 4MB pages.
6. [C8, 3 pts] A computer system is experiencing a very low CPU utilization but having a very busy paging disk. Answer the following:
    a. What would be the cause of this problem?
    b. Can killing some processes help improve the CPU utilization? Why?
    c. Continued from above, how about raising the CPU clock rate?
7. [C9, 6 pts] Consider two programs: one program that accesses random location of a large, sparse matrix in the memory and the other program that sequentially searches a large array for a pattern match. Which one(s) of the programs can substantially benefit from pre-paging? Why?
8. [C9, 6 pts] LRU (Least-Recently Used) and LFU (Least-Frequently Used) are two representative page replacement algorithms. Now consider the memory footprint of a process (shown below). Which one of LRU and LFU will be a better choice for the memory workload? Why?



X-axis: time
Y-axis: page number

9. [C9, 8 pts] There are two ways to duplicate files
   a. Mapping the source and destination files using mmap() and then using memcpy() to copy data
   b. Opening the two files using fopen() and then using fread()/fwrite() to copy data.

   Now you are duplicating large files on a fast storage device, e.g., an NVMe SSD. Which one of the above will be faster? Why? Give at least two major reasons.

10. [C10, 4 pts] Consider hard links in file systems. What is the dangling pointer problem? How is it resolved in common file systems?

11. [C10, 6 pts] There are three basic steps to prepare a hard drive for file storage: partitioning, formatting, and mounting. Describe the meaning of each step and why they are necessary.

12. [C11, 8 pts] There are two types of storage devices: fixed and removable. The former refers to regular hard drives and SSDs, while the latter can be USB thumb drives, CD-ROMs, and floppy discs. In addition, removable drives can be unplugged by the user unexpectedly. In consideration of this, which one of the disk-cache write policies, write-back and write-through, is more suitable for removable drives? Why?

13. [C11, 6 pts] WAL (Write-Ahead Logging) is a common technique to implement file-system journaling. It first collects a set of self-contained block writes into a transaction, writes the transaction to a reserved log space on the disk, and then applies the block writes to the file system. Now consider the following power fail scenarios. Explain how the file system can easily recover from these events without having any structural inconsistency:
   a. Power goes off when the file system is writing a transaction to the log
   b. Power goes off when the file system is done logging the transaction and applying the writes to the file system

14. [C11, 6 pts] Explain how the following techniques improves file system performance
   a. Having a few direct pointers in an inode
   b. Embedding small files in directories
   c. Dividing the disk space into cluster groups

15. [C12, 8 pts] SSTF (Shortest-Seek Time First) is an intuitive, efficient disk scheduling algorithm. However, it is rarely employed in real operating systems. Discuss the critical problem of SSTF and also propose an enhancement to deal with it.

16. [C12, 6 pts] RAID-5 is a commonly used disk array design. Answer the following:
   a. What is the advantage of RAID-5 compared with RAID-0?
   b. What is the advantage of RAID-5 compared with RAID-4?

17. [C12, 6 pts] Garbage collection in Solid-State Disks (SSDs) involves a series of page reads, page writes, and block erasure operations. Answer the following:
   a. Why is garbage collection necessary in flash-based SSD?
   b. A garbage-collection policy decides which flash block will be erased. What is the primary performance consideration of its decision making?

****You must justify your answer with sufficient details to get full marks. (total 105)*