Introduction to Operating Systems, 2024 Fall, NYCU CS
Question Sheet for Final

Part A: Basic Problems [total:100pts]

1. [7pts]Answer the following regarding dynamic memory allocation
   a. [2pts]What is internal fragmentation?
   b. [2pts]What is external fragmentation?
   c. [3pts]What kind(s) of fragmentation the Buddy System is subject to? Explain your answer.
2. [9pts]Consider a CPU architecture, for which the process virtual address space is addressed by 32 bits and the physical memory space is addressed using 34 bits. The page size is 2048 bytes. The CPU employs two-level paging, and the first-level page table has 512 entries. Answer how many bits are required by the following (ignore control bits):
   a. [3pts]A page number
   b. [3pts]A frame number
   c. [3pts]A second-level page table entry
3. [4pts]Regarding inverted page tables
   a. [2pts]What is their main benefit in terms of space overheads compared with conventional (forward) page tables?
   b. [2pts]Propose a method for inverted page tables to support shared memory.
4. [8pts]Explain the meanings and applications of the following control bits for virtual memory management:
   a. [2pts]Valid/invalid bit
   b. [2pts]Clean/Dirty bit
   c. [2pts]Reference bit
   d. [2pts]Read-Only/Read-Write bit
5. [3pts(1+2)]See Fig. 1. What is the expected performance curve of page faults under different physical memory sizes? Explain your answer.

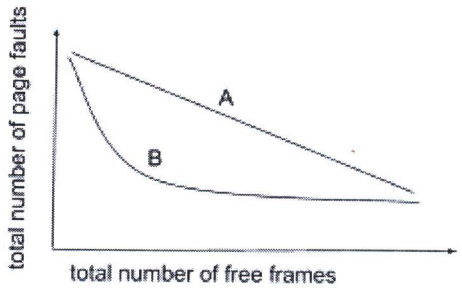

Fig. 1

6. [7pts]LRU and LFU are representative page replacement algorithms. Answer the following:
   a. [4pts(2+2)]What are the major drawbacks of LRU and LFU?
   b. [3pts]It is infeasible to implement LRU without simplification in the virtual memory system. Instead, we use LRU approximations such as NRU (e.g., the Clock algorithm). Why the original LRU is impractical in the memory system?
7. [6pts]Explain how the following designs improve the performance of virtual memory:
   a. [2pts]Background flushing (e.g., Linux bdi_writeback)
   b. [2pts]Background reclaiming (e.g., Linux kswapd)
   c. [2pts]Prefetching
8. [9pts]Answer the following regarding file links
   a. [2pts]What is a soft link?
   b. [2pts]What is a hard link?
   c. [3pts(1+2)]Which one(s) cannot span over different file systems? Why?
   d. [2pts]How Linux handles the dangling problem for the two types of links?
9. [3pts]Describe the purposes of the following Linux VFS objects
   a. [1pts]inodes
   b. [1pts]File objects
   c. [1pts]Dentry objects
10. [6pts]Extent allocation pre-allocates contiguous, free disk space for files. Answer the following
    a. [2pts]Why does extent allocation benefit file system performance when the storage device is HDD?
    b. [2pts]Continued from question a, is it still valuable for SSD?
    c. [2pts]What is the drawback of extent allocation in terms of disk space utilization?

11. [10pts]Ext4 file system is a journaling file system. It guarantees the atomicity of file system transactions. Answer the following:

a. [2pts]What is a file system transaction?
b. [2pts]What does transaction atomicity mean?
c. [3pts]Describe the whole process of Ext4 journaling, beginning with the arrival of incoming data to the WAL journal and continuing until the data are written into the file system.
d. [3pts]Continued from c, how does the system recover from a power interruption if a transaction is partially written to the WAL journal?

12. [10pts]Answer the following questions regarding log-structured file systems:

a. [4pts(2+2)]What is out-of-place updating? How does it optimize random write performance?
b. [3pts]Why cleaning (compaction or garbage collection) is necessary?
c. [2pts]The performance of a log-structured file system significantly degrades when the file system is nearly full. Why?
d. [1pts]Name a real log-structured file system.

13. [6pts]Consider a disk drive with 5000 cylinders numbered from 0 to 4999. The disk arm is currently positioned at cylinder 143, with the previous request being served at cylinder 125. The queue of pending requests, listed in FIFO order, is as follows:

86, 1470, 913, 1774, 948, 1509, 1022, 1750, 130

Calculate the total distance traveled by the disk arm (in terms of cylinders moved) to satisfy all pending requests using each of the following disk-scheduling algorithms. You *must* also specify the service order of requests.

a. [3pts]SSTF
b. [3pts]SCAN

14. [6pts]Answer the following questions regarding RAID-5:

a. [2pts]What is a parity block?
b. [2pts]Why are parity blocks evenly distributed over all disks?
c. [2pts]How does RAID-5 tolerate a single-disk failure?

15. [6pts]Answer the following questions regarding SSD:

a. [2pts]Why is out-of-place updating necessary for efficient writes in flash memory?
b. [2pts]What is garbage collection and why it is necessary in flash memory?
c. [2pts]The write amplification factor (WAF) is defined as follows

$$WAF = \frac{Total\ amount\ of\ data\ written\ to\ flash}{Total\ amount\ of\ data\ written\ by\ the\ host}$$

For example, the host generates 100MB of data to the SSD, but the internal flash memory receives 300MB of writes. The ideal WAF is 1.0. However, for consumer-grade SSDs, their WAF is generally between 3 and 7. Explain why.

Part B: Advanced Problems [total: 6pts]

16. [3pts]Design a memory allocation and de-allocation sequence, for which Best Fit fails to service an allocation request while there is sufficient memory and the space utilization is 1% only.

17. [3pts]Design a case in which SCAN outperforms SSTF, and another case for the opposite.

Total: 106 pts

You must sufficiently explain your answer for each question to get full marks.
You are solely responsible for your answer. TAs provide explanations just for good intentions.