CS323 – Final Preparation Exercises

23 May 2019

Question 1: Disk optimization

Two of the primary disk optimizations are disk scheduling and clever disk allocation. One is known to be more effective under high load, and the other is known to be more effective under low load. Which one is which and why? Structure your answer as follows:

1. Disk scheduling is more effective under low load / high load (pick one).
2. Reason for answer to 1
3. Clever disk allocation is more effective under low load / high load (pick one).
4. Reason for answer to 3

Question 2: Disk Scheduling

Suppose that a disk drive has 5000 cylinders, numbered from 0 to 4999. The drive is currently serving a request at cylinder 143, and the previous request was at cylinder 125. The queue of pending requests is:
86, 1470, 913, 1774, 948, 1509, 1022, 1750, 130

Starting from the current head position, what is the total distance (in cylinders) that the disk arm moves to satisfy all the pending requests for each of the following disk-scheduling algorithms?

1. FCFS
2. SSTF
3. SCAN
4. C- SCAN

Question 3: In-memory and on disk data structures

	Active File Table	**Open File Table**
Scope		
Each entry corresponds to		
Each entry contains		

A. Fill in the following table, comparing the Active File Table and the Open File Table.

B. List the steps that take place when a process invokes *tid = Open(uid).* In particular, show any accesses to disk and any updates to the above data structures. Assume that the file with identifier *uid* exists, but that all in-memory data structures are empty when the operation is invoked.

Question 4: Log-structured Key-Value Stores

Similar principles that are used by the log-structured file system can be applied to different contexts. One such example is *log-structured merge key-value stores (LSMs)*. LSMs are data stores holding key-value pairs. They have a *sorted* in-memory component and an on-disk component. All writes are performed (only) on the in-memory component. Once the in-memory component gets full, it is sequentially written to the disk component. The disk component hence contains multiple files corresponding to the former memory components, which were persisted to disk.

A. Describe what a read operation look likes in this LSM. In other words, given a particular key, how do you find the corresponding value.

B. A problem of the described LSM is that the files stored on disk could contain many duplicates of the same key, wasting space. Describe a “cleanup” operation that would solve this problem.

C. For what kinds of workloads or applications would the LSM described above be the most suitable for? Motivate your answer.

Question 5: RAID

Consider a 4-disks, 256GB-per-disks RAID array. What is the available data storage capacity if:

A. Disks are organized as RAID 0?

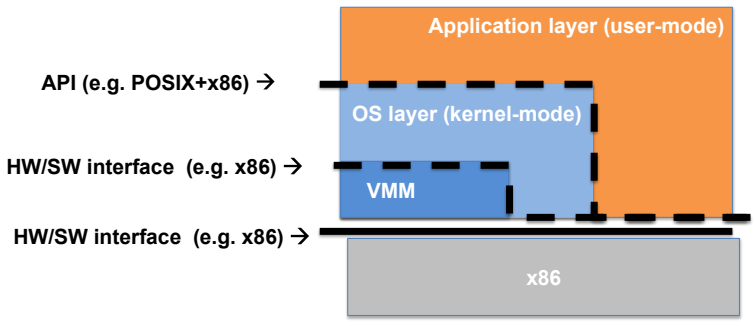
B. Disks are organized as RAID 5?

Question 6: SSDs and LFS

What makes a log-structured file system suitable as a file system for SSDs?

Question 7: Virtualization

Assume an architecture that meets the Popek/Goldberg criteria for virtualization and follows the layering of this figure.



Provide one example of a transition across each of the 5 boundaries that are pictorially represented in the figure, i.e.

A. Application layer – x86

B. Application – OS

C. OS – x86

D. OS – VMM

E. VMM – x86