

**FIRST PART OF THE FINAL AND  
GRADUATE PART OF THE QUALIFYING EXAMINATION**

This exam is closed book. You can have two sheets of self-prepared notes. UH expels cheaters.

**1. *Disk Internal Buffers:***

Nearly all disk drives have now large internal buffer that acts both as a speed-matching buffer and as a block cache. Lower memory prizes have resulted in a dramatic increase of the sizes of these internal buffers, with sizes between 2 and 16 Megabytes being typical.

A recent study seems to indicate that there is little benefit in increasing the size of these buffers, pointing out that “all modern operating systems already use large I/O buffer caches for their read and writes.”

Assuming that both buffers are managed by Least-Recently-Used policies, explain how the existence of a large I/O buffer cache in main memory can affect the caching performance on the disk drive internal buffer. Are there better ways to manage these internal buffers or should we conclude that large internal buffers on the disk drive are a bad idea? (20 points)

**2. *Serverless File Systems Using Personal Workstations***

Several current projects aim at using the vast storage and processing resources of personal workstations to build serverless file systems. Instead of storing user files on dedicated machines, these proposals distribute them among the workstations of other users.

These proposals raise interesting questions in terms of file availability (users will keep the right to turn off their workstations whenever it pleases them) and security (we must protect files against the possibility of *malicious actions* by the owners of the workstations storing other people files). How would you address these issues? (20 points)

**Hint:** Malicious users will always constitute a small minority of the total user population.

1. RAID-2 and RAID-4 disk organizations have never been used in any commercial product. Explain why? (2×5 points)
2. What is the major advantage of Coda *callbacks*? What is its major disadvantage? What can we do to minimize this disadvantage? (3×5 points)
3. Consider a diskless client trying to access a file named “/usr/joe/6360/paper.doc” that is stored on its NFS server.
  - a) Assuming that the client already has a handle for its root directory, how many *lookup()* requests will it issue? (5 points)  

\_\_\_\_\_ requests
  - b) What does NFS do to speed up these requests? (5 points)
  - c) What will be the results of the last lookup in the client chain of requests? (5 points)
4. What criterion does Elephant use to decide that a specific version of a file should be kept forever? (5 points)
5. What are the main advantage and the main disadvantage of “*stateful*” servers? (2×5 points)
6. What makes Kerberos poorly suited to environments where users have to log in from workstations that cannot be trusted? (5 points)