

Name: _____

Instructions

There are four problems on the exam. You may find questions that could have several answers and require an explanation or a justification. As we've said, many answers in storage systems are "It depends!". In these cases, we are more interested in your justification, so make sure you're clear. Good luck!

Problem 1 : Short answer. [56 points]

- (a) A UNIX system administrator runs a web server on a machine that gets all of its files from an NFS server. Using a different client machine, the administrator removed permissions for a number of the web server files, and doing so caused the web server to crash. The administrator is confused, because the exact same permission removal did not cause a crash the previous year, when the web server's files were all stored in a local file system on the same machine as the web server. Briefly explain why the use of NFS storage (instead of the local file system) could result in such a problem.
- (b) The basic SCSI bus has 8 data lines and can support up to 8 devices. Why can it not support more?

- (c) A storage system engineer noticed something interesting when experimenting with different disk array configurations: RAID level 5 performs better than RAID level 4 (given the same set of disks) for a workload consisting entirely of small random reads. Briefly explain why.
- (d) In most parallel file systems, reliability is provided by RAID “underneath” (i.e., within) each server. Why does this approach require expensive dual-attached RAID controllers to provide high reliability?
- (e) A system administrator has 100 disks at his disposal. He is considering configuring them as either ten 10-disk RAID-5 arrays or one 100-disk RAID-5 array. Quantify the improved reliability that would come from the configuration with 10× more capacity being used for redundancy. (Assume a MTBF of 1000 days for each disk.)

(f) Joe uses a high-end storage server that includes RAID support for reliability, LFS-based mechanisms for improved write performance, and snapshots for on-line backup. Unfortunately, he has almost completely filled the storage capacity. To fix the problem, he identified and deleted a number of large, old files. But, when he checked the capacity utilization right after deleting the files, he found that there was no additional free space. Briefly explain why.

(g) Janet needs to access some files in an AFS directory for her work. After authenticating properly to her client machine and to Kerberos, she finds that the system allows her to access some of the files in the directory and not others. Give one possible cause for this behavior.

Problem 2 : Disk Arrays. [20 points]

You are given a disk array with 5 disks. Assume that each disk has an average 6ms seek time, a 6ms full revolution time (i.e., 6ms to rotate 360 degrees), synchronized spindles, and a transfer rate of 100MB/s. The array is configured to use RAID-4.

(a) Assume a workload consisting of random aligned 100KB write requests. Compute the average request service time and the maximum array throughput for each of these stripe unit size options:

- 100KB.

- 25KB.

(b) Suggest one simple configuration change (with no new hardware) that could significantly improve the throughput of the 100KB case.

Problem 4 : Bonus Questions. [Maximum of three points]

(a) Which instructor will be moving to the new Gates building in August?

(b) Which instructor watches University of Michigan football games with his sons?

(c) What kind of camp (i.e., what theme) should Prof. Ganger's kids attend this summer?

(d) What color was George Washington's white horse?