1. **Virtual Memory.**

In this question, we'll see how long a simple load instruction takes to execute, given some different assumptions about the virtual memory system.

- Assume we have a linear page table per process, but the system we are using has no TLB and no swapping to disk. Assume each memory access takes M units of time. How long does it take to execute a single load instruction (e.g., `mov VirtualAddress, register`), in terms of memory accesses?

2 — 4M — 2M => X (but OK on rest)

- Now assume the same system, but with a TLB (still no swapping). Assume memory access is the dominant cost, and it takes M units of time to access memory. What is the fastest (i.e., best case) the load instruction above will execute, assuming TLB hits?

2 — 2M or 0M (w/ h/w cache)

- Assume now we have a two-level page table with a page directory. Assume each reference to the TLB is a miss, but that all referenced pages are found in memory (no swapping again). What is the slowest (i.e., worst case) time for the load instruction above?

3 — 6M

- Assume the same system again, but this time the memory location(s) referenced by the load instruction also might be on disk (due to swapping). Disk accesses (in this simple model) take D time units. What is the worst case time for the instruction to execute?

3 — worst: page dir, page table page, page (inst)
" " " (data)

(assume PT in mem)
swap out; swap in => 4M + 4D

2. **Concurrency.**

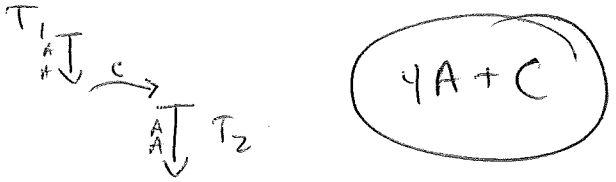
Given a basic spin lock, assume that locking the spin lock takes A time units (if no one is holding the lock); unlock also takes A time units. Assume further that a context switch takes C time units, and that a time slice is T time units long.

Assume this code sequence, executed by two threads on one processor at roughly the same time:

```
mutex_lock();
do_something(); // takes no time to execute
mutex_unlock();
exit();         // takes no time to execute
```

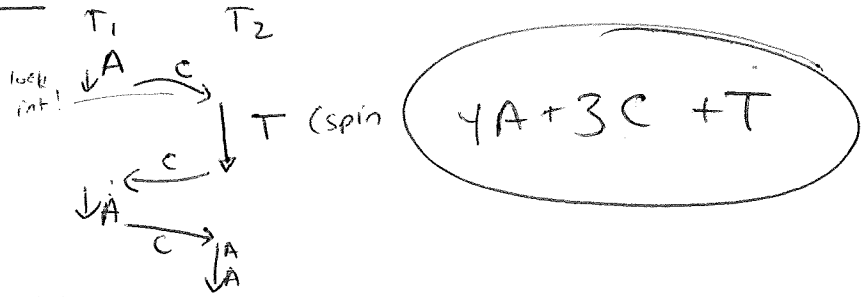
- What is the best-case time for the two threads on one CPU to finish this code sequence?

3



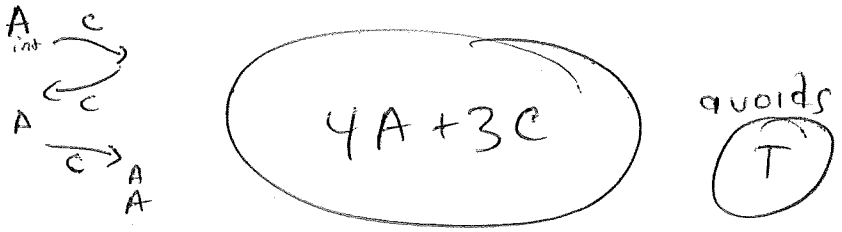
- What is the worst-case time for the two threads to finish this code sequence? Assume that only three context switches can occur at a maximum.

3



- If the spin lock is instead changed to a queue-based lock, how does that change the worst-case time?

3



(1 free)

3. **Disks.**

In this question, we'll perform some simple calculations on a highly-simplified disk.

- Assume a simple disk that has only a single track, and a simple FIFO scheduling policy. The rotational delay on this disk is R; there is no seek cost (only one track!), and transfer time is so fast we just consider it to be free. What is the (approximate) worst case execution time for three (3) requests (to different blocks)?

2

~ 3R

- Now assume that a shortest-access-time-first (SATF) scheduler is being used by the disk (but it still only has a single track). What is the worst-case time for three requests (to different blocks) now?

2

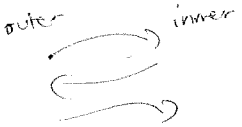
~ R

- Now assume the disk has three tracks. The time to seek between two adjacent tracks is S; it takes twice that to seek across two tracks (e.g., from the outer to the inner track). Given a FIFO scheduler, what is the worst-case time for three requests?

3

single request worst: 2S + R

3 reqs => 6S + 3R



- Same question, but for a SATF scheduler.

3

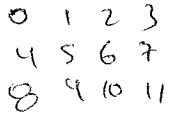
should be able to do in roughly one sweep across disk (but might have to wait at each stop)

~ 2S + 3R

4. **RAIDs.**

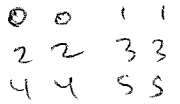
For this question, we'll examine how long it takes to perform a small workload consisting of 12 writes to random locations within a RAID. Assume that these random writes are spread "evenly" across the disks of the RAID. To begin with, assume a simple disk model where each read or write takes D time units.

- Assume we have a 4-disk RAID-0 (striping). How long does it take to complete the 12 writes?



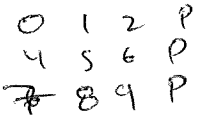
~3D

- How long on a 4-disk RAID-1 (mirroring)?



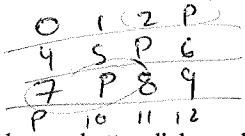
~6D

- How long on a 4-disk RAID-4 (parity)?



~12D ×2 ⇒ 24D

- How long on a 4-disk RAID-5 (rotated parity)?



~6D ×2 ⇒ 12D

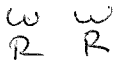
- Now assume we have a better disk model, in which it takes S time units to perform a random seek and R units of time to perform a full rotation; assume transfer is free. How long do the 12 random writes take to complete on a 4-disk RAID-0?

$\sim 3(S + \frac{1}{2}R)$

- How long on a 4-disk RAID-1 (mirroring)?

$\sim 6(S + \frac{R}{2})$ ← but really slightly longer as you get worst case of 2 seeks, rotates

- How long on a 4-disk RAID-4 (parity)?



1 req: 2R ⇒ $S + \frac{R}{2}$
2R ⇒ R

$\sim 12(S + \frac{3R}{2})$

- How long on a 4-disk RAID-5 (rotated parity)?

$\sim 6(S + \frac{3R}{2})$

5. **A Basic File System.**

For this question about basic file systems, assume a simple disk model where each disk read of a block takes D time units. Also assume the basic layout is quite like the very simple file system or FFS.

- Assume that all data and metadata begin on disk. Assume further that all inodes are in separate blocks, and that each directory is only one block in size. How long does it take to open the file `/a/b/c/d.txt`?

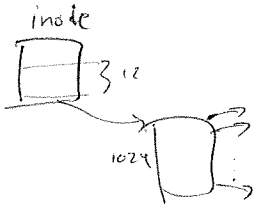
read root inode
root data
a inode
a data
b inode
b data
c inode
c data
d.txt inode

3

~ 9D

- Assume after opening the file, we read the file in its entirety. It is a big file, containing 1036 blocks. The inode itself has room for 12 direct pointers and 1 indirect pointer. Disk addresses are 4 bytes long, and disk blocks are 4KB in size. After opening it, how long does it take to read the entire file?

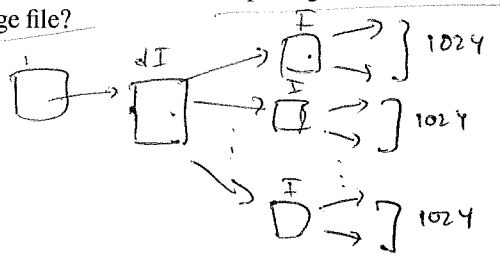
~~3~~



1037 D

- Now assume a new inode structure is introduced, in which there is only one pointer: a double indirect pointer, which points to the double indirect block, which can point to 1024 indirect blocks, each of which can point to 1024 blocks. After opening the file, how long does it take to perform 50 random reads within a very large file?

3



D (double indirect)
+ 2D (for ind/data)
x 50 of each block)

⇒ 101 D

- How long does it take to close a file, approximately (assuming disk accesses are the dominant cost)?

~~3~~ 1

~ 0D

6. **Journaling File Systems.**

For this question about journaling file systems, assume the following disk model: it takes S time units to seek to any location on the disk; the full rotational delay is R time units; transferring data takes T units, regardless of the amount of data transferred.

- Assume we have a data journaling file system in which all data and metadata are logged before being written to disk. How long does it take to append a block to an existing file, assuming that the all relevant metadata and data was in the OS page cache to begin with (i.e., no reads from disk need occur)?

2

journal writes: $S + \frac{R}{2} + T + R + T \Rightarrow S + \frac{3R}{2} + 2T$

(T_B + I, B, D; T_E)

+

checkpoint: $3\left(S + \frac{R}{2} + T\right) \Rightarrow 3S + \frac{3R}{2} + 3T$

- Now assume we have journaling, but only for metadata; data blocks are written directly to their final location. How long does the append take in this case?

3

data first (D): $S + \frac{R}{2} + T$ +

journal (I, B): $S + \frac{3R}{2} + 2T$ +

checkpoint (I, B): $2\left(S + \frac{R}{2} + T\right) \Rightarrow 2S + R + 2T$

- Now do the same calculation, but without journaling at all. How long does the append take without journaling?

2

$\Rightarrow 3\left(S + \frac{R}{2} + T\right)$

write I, D, B

- Finally, let's assume we're using full data journaling again. How long does it take to create a new empty file? Again assume that all relevant structures are in cache to begin with.

3

create: dir I, dir data, new I, Inode bitmap

journal: $S + \frac{3R}{2} + 2T$ +

checkpt: $4\left(S + \frac{R}{2} + T\right) \Rightarrow 4S + 2R + 4T$

1R 2S 2S ~ 3R+2S
~360 80 80

7. **LFS, the Log-structured File System.**

We'll now analyze the performance of LFS, the log-structured file system. Assume it takes S time units to seek to a segment, and T time units to read or write it after that seek. If, however, we access segments in sequential order (segment 0, then 1, then 2, ...), no seek cost is incurred.

- Assume we have a newly-created empty file system. We now write out a file that fills 100 segments. How long does it take to complete this write sequence?

2 S + 100 T

- Assume we now read this file (and that the reads do not hit in the memory cache); how long does it take to read the entire file?

2 S + 100 T

- Assume we now read the file backwards, one segment at a time (and that the reads do not hit in the memory cache); how long does this backwards read take?

2 100 S + 100 T

- Assume we have an old file system, in which many segments have been repeatedly cleaned, leaving a few hundred free segments scattered over the disk. How long does it take to write out 100 segments on this file system?

2 100S + 100 T

- Assume now that the disk is nearly full, and that to write a segment, we have to clean two segments (which produces one compact segment, and one free one). In this case, how long does it take to write out 100 segments?

2

clean two: 2S + 2T] read two old
S + T] write compact
S + T] write new
} 100 times

~ 400 S + 400 T

8. **NFS, Sun's Network File System.**

We'll now model the time of certain operations in the Sun Network File System. The only costs to worry about are network costs; assume everything else is free. Assume any "small" message takes S units of time to be sent from one machine to another, whereas a "bigger" message (e.g., the size of a disk block, 4KB) takes B time units. If a message is larger than 4KB (e.g., 8KB), it should take proportionally longer (e.g., $2B$ for 8KB).

- How long does it take to open a 100-block (400 KB) file called `/a/b/c.txt` for the first time? (assume that the root directory is indeed the root of the NFS file system)

2

lookup(root fd, "a") → S
lookup(a fd, "b") → S
lookup(b fd, "c.txt") → S

≈ 6S

- How long does it take to read the file?

2

100 × (S + B) ⇒ 100S + 100B

- How long does it take to re-read that file immediately?

2

≈ 0

- How long does it take to re-read that file after a little while?

2

check if changed ≈ 2S

re-read from cache : 0

- How long does it take to re-read that file a long time later, after many other memory-consuming programs have run on the client?

2

not in mem any more

⇒ ≈ 100(S + B)

9. **AFS, the Andrew File System.**

We'll now model the time of certain operations in the Andrew File System. The only costs to worry about are network costs; assume everything else is free. Assume any "small" message takes S units of time to be sent from one machine to another, whereas a "bigger" message (e.g., the size of a disk block, 4KB) takes B time units. If a message is larger than 4KB (e.g., 8KB), it should take proportionally longer (e.g., $2B$ for 8KB).

- How long does it take to open a 100-block (400 KB) file called `/a/b/c.txt` for the first time? (assume that the root directory is indeed the root of the AFS file system)

2 6S (lookups) + ~~[illegible]~~ S (fetch) + 100B (data)

- How long does it take to read the file?

all local: ~0

2

- How long does it take to re-read that file immediately?

2 ~0

- How long does it take to re-read that file after a little while?

2 ~0

- How long does it take to re-read that file a long time later, after many other memory-consuming programs have run on the client?

2 ~0

10. **Virtual Machine Monitors.**

With virtual machine monitors (VMMs), a lot of excess overhead arises due to the extra level of virtualization beneath the OS. In this question, we'll examine these overheads. Assume the only costs we are concerned with are T, the cost of trapping into the VMM, and R, the cost of returning from such a trap.

correction

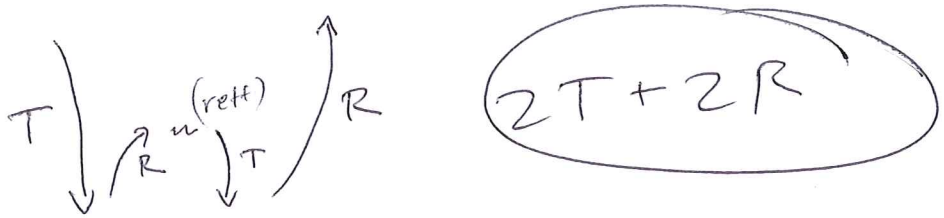
priv. entry

- When performing a system call on an OS *not* running on a VMM, how long does it take in terms of T and R?

2

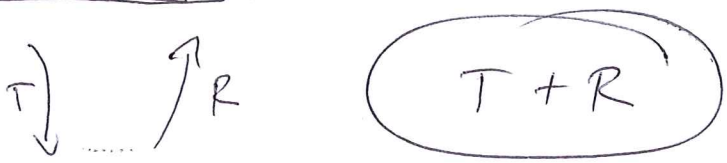
- When performing a system call on an OS running on a VMM, how long does it take in terms of T and R? (assume the system call is a simple one, such as getpid())

2



- When executing an instruction that causes a TLB miss (assuming a software-managed TLB), how long does it take on an OS *not* running on a VMM?

3



- When executing an instruction that causes a TLB miss (assuming a software-managed TLB), how long does it take on an OS running on a VMM?

3