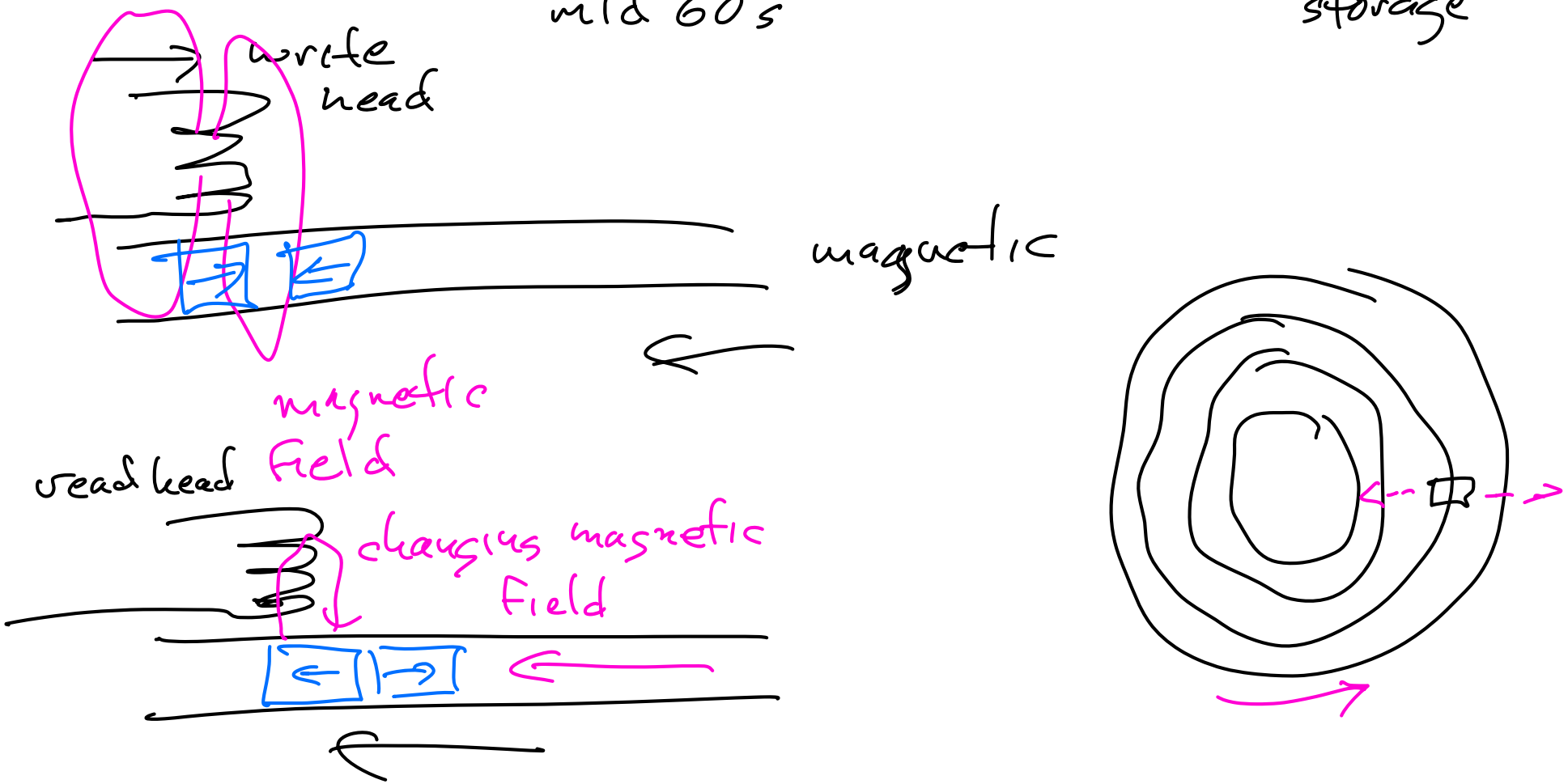
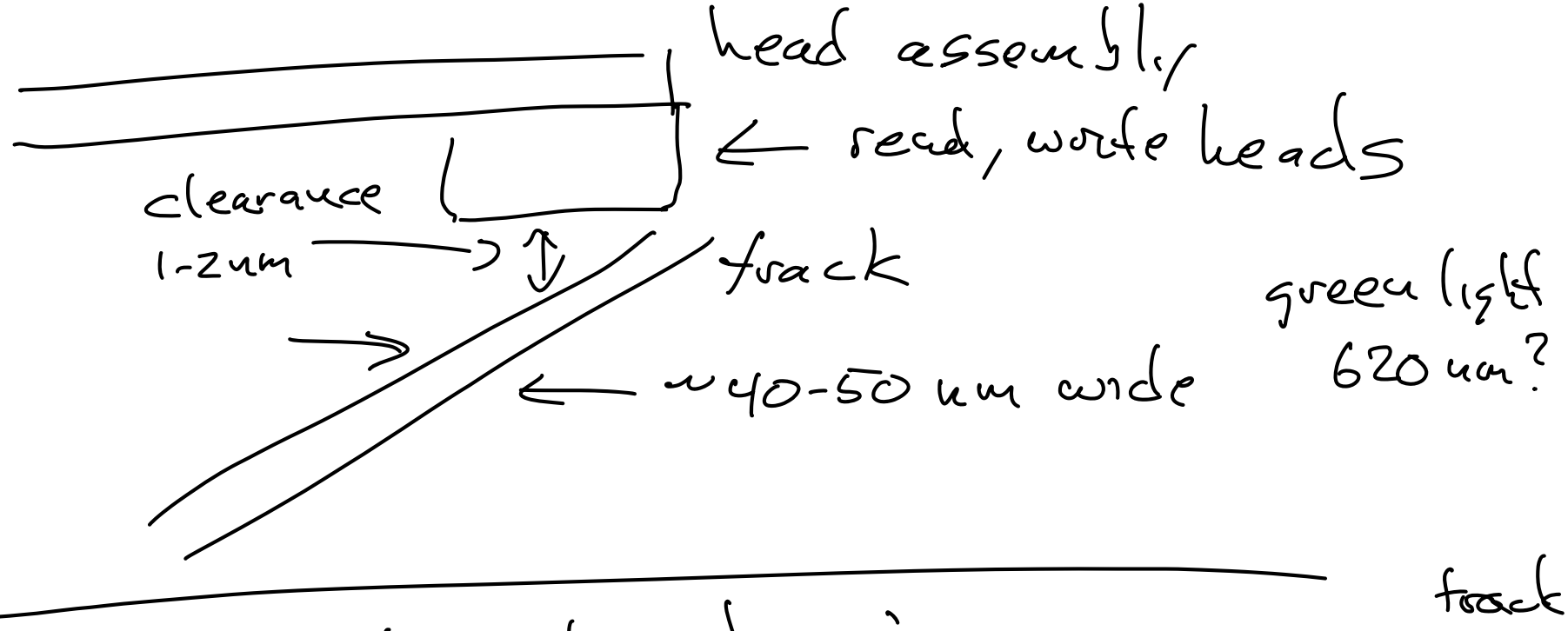
CS 3650 – Computer Systems
Spring 2024
Peter Desnoyers

Lecture 19, Tue Mar 19, 2024

File systems, disks & SSDs, Lab 4

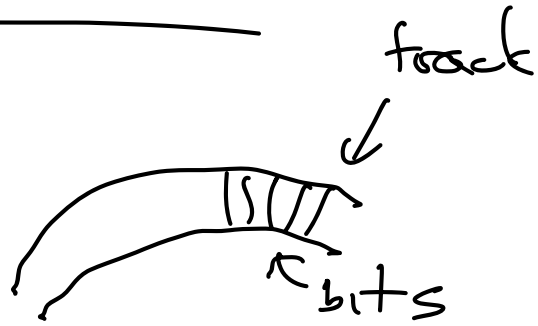
Hard drive - 1959 (58?) — 2015 primary storage
mid 60s



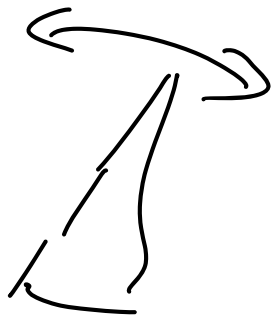


Hard drive characteristics:

- rotation speed 7200 RPM
- transfer rate - - -200MB/s
- seek time
  15 ms max



200-250 MB/s



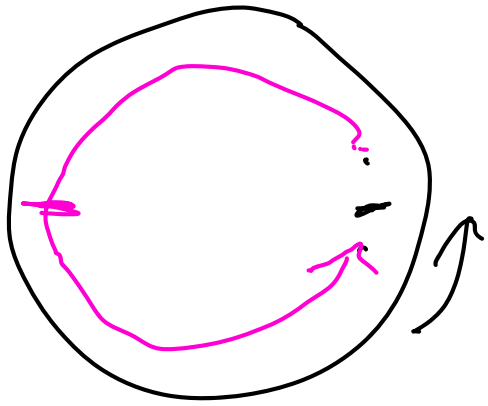
transfer time for random I/O:

seek time + (½ max = average)

½ rotation (average) +

transfer time

7200 RPM
= 8.3 ms



random 4KB block:

seek: 7.5 ms

rotation: 4.2 ms

fixed ≈ 20 MB

variable (e.g. per-byte)

4KB transfer: 20 µs = 0.020 ms

200 MB/s

≈ 11.7 ms

random 1MB:

transfer: 5 ms

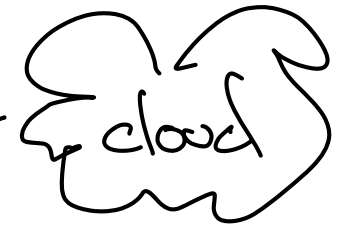
10 MB: 11.7 overhead

50 ms data transfer

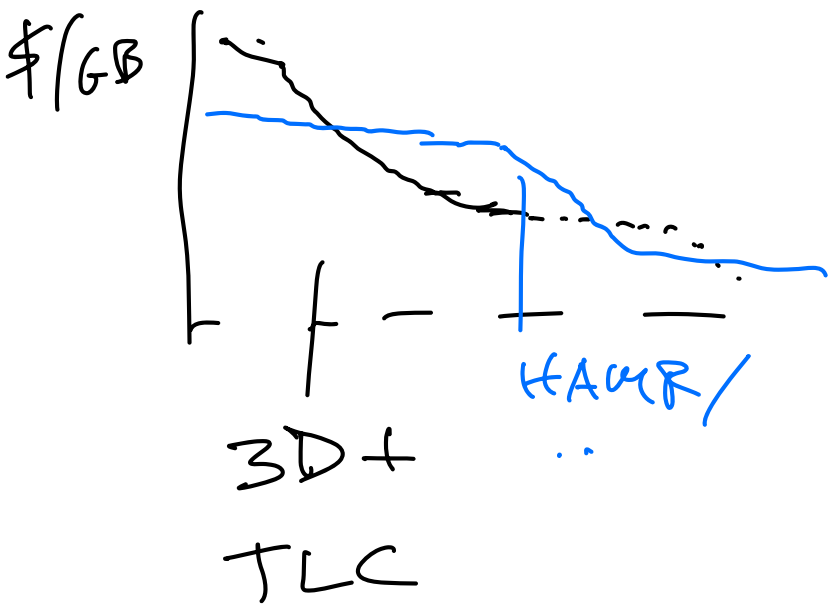
Are hard drives dead?

bulk storage

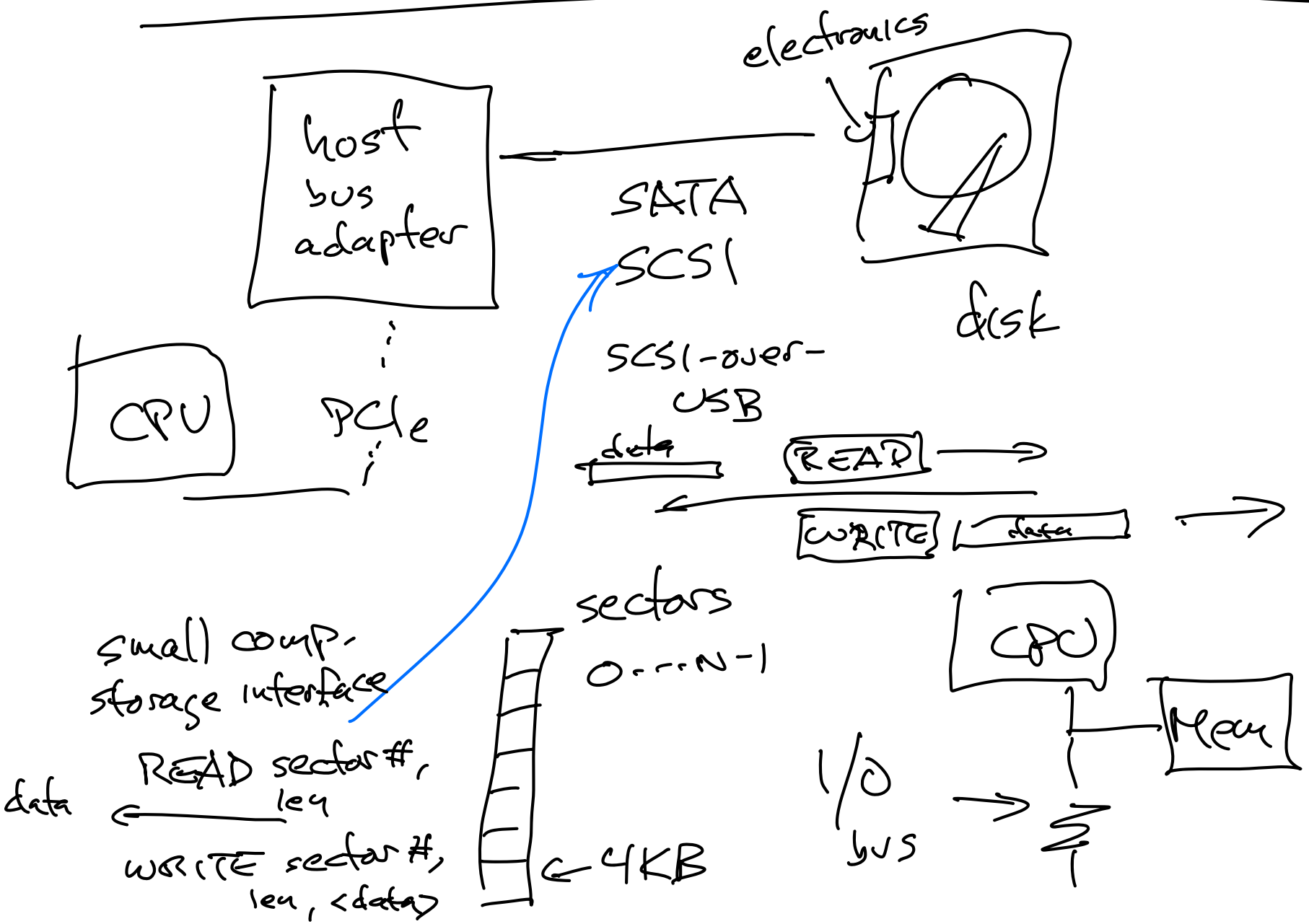
cheap, performance not critical



backup drives



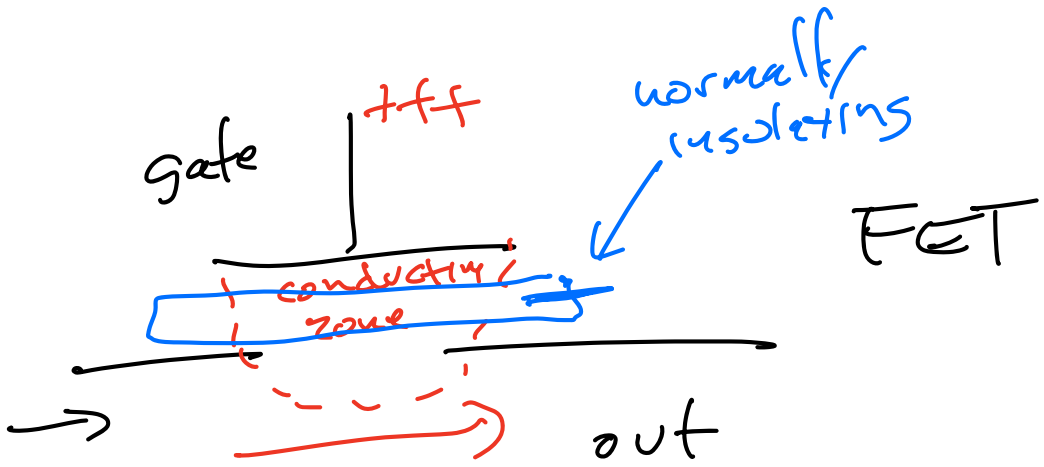
How your computer talks to a disk



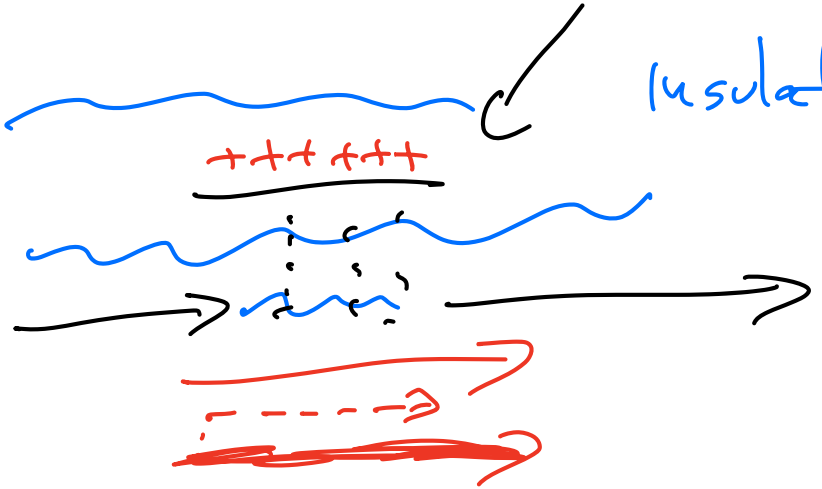
SSDs

NAND Flash

DRAM: + capacitor



"floating gate"



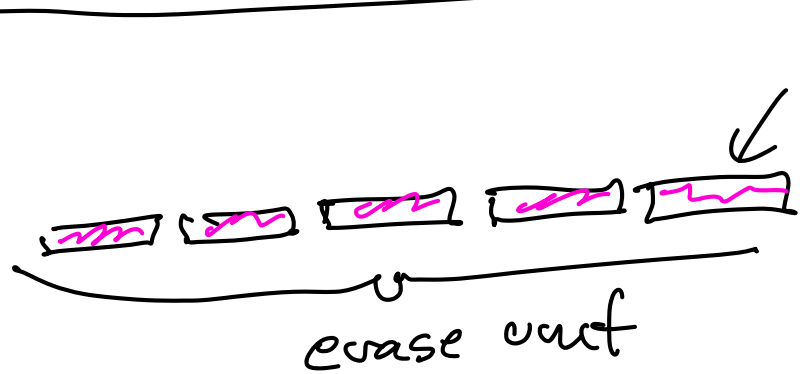
a lot: 1 1

more: 1 0

a little: 0 1

none: 0 0

Flash translation layer & garbage collection



write once (block)

read (block)

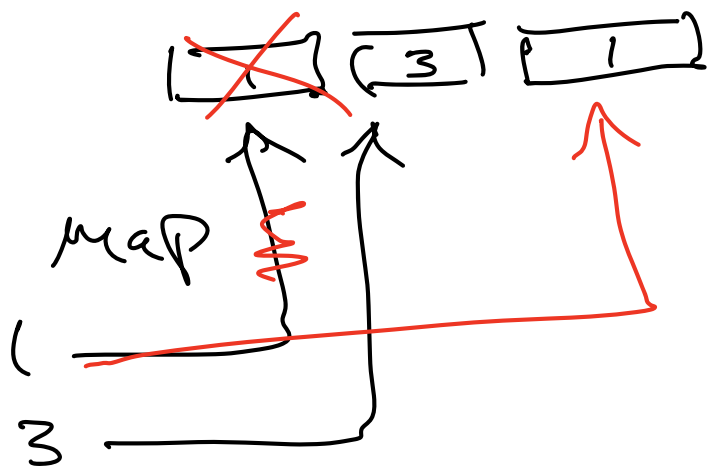
erase (erase unit)

block write, bulk erase

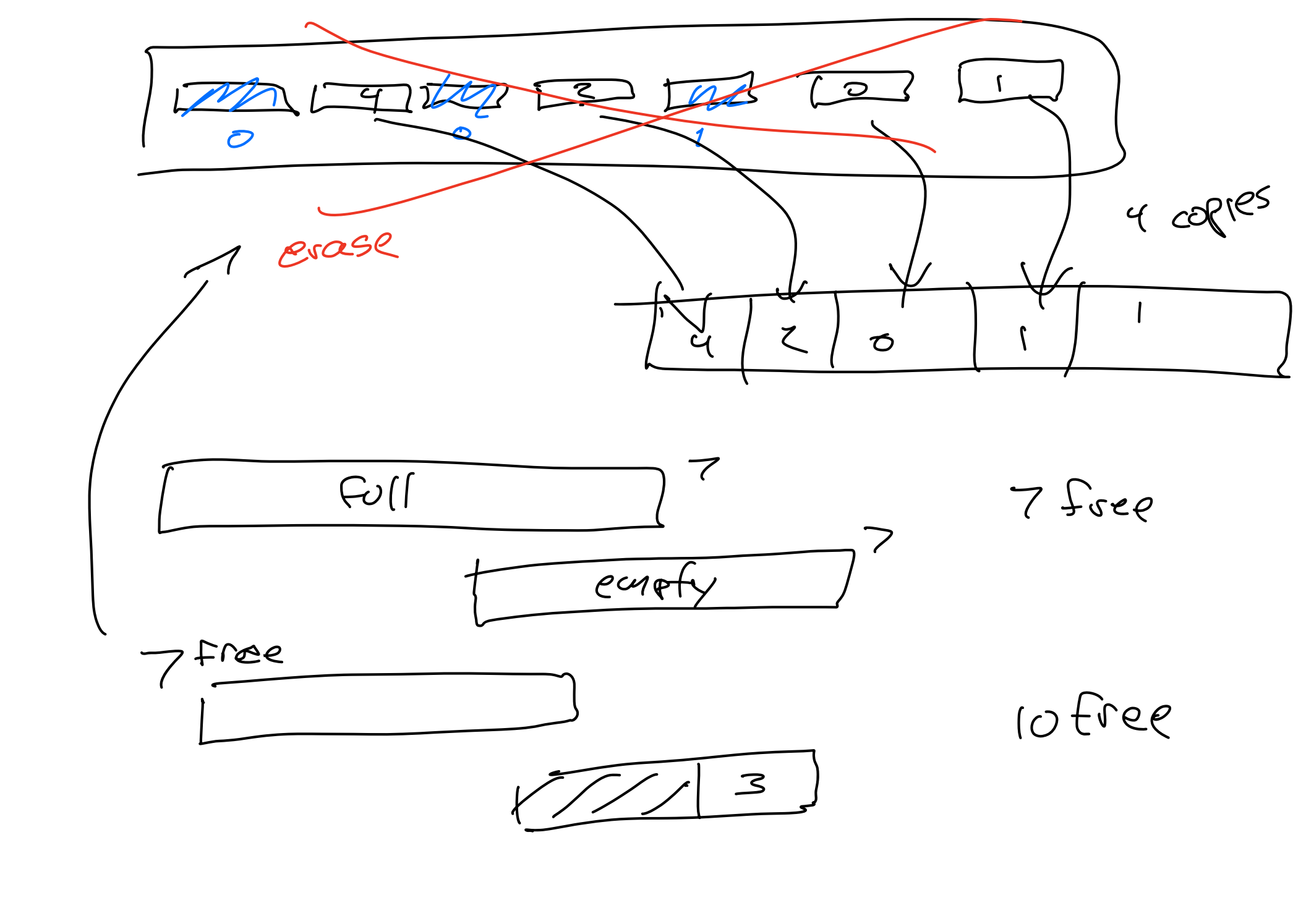
→ write 1

→ write 3

→ write 1.



out-of-place write


4
3
0
1
0
0
1
erase
4 copies
4
2
0
1
1
full
7
7 free
empty
7
7 free
3
10 free

Write amplification:

internal flash write ops =

"real" writes (from computer)

+ GC writes (moving data so can erase a unit)

higher write amp. if:

- random small writes
- low free space %

TRIM → tell SSD to forget blocks not used by file system (discard)

Lab 4

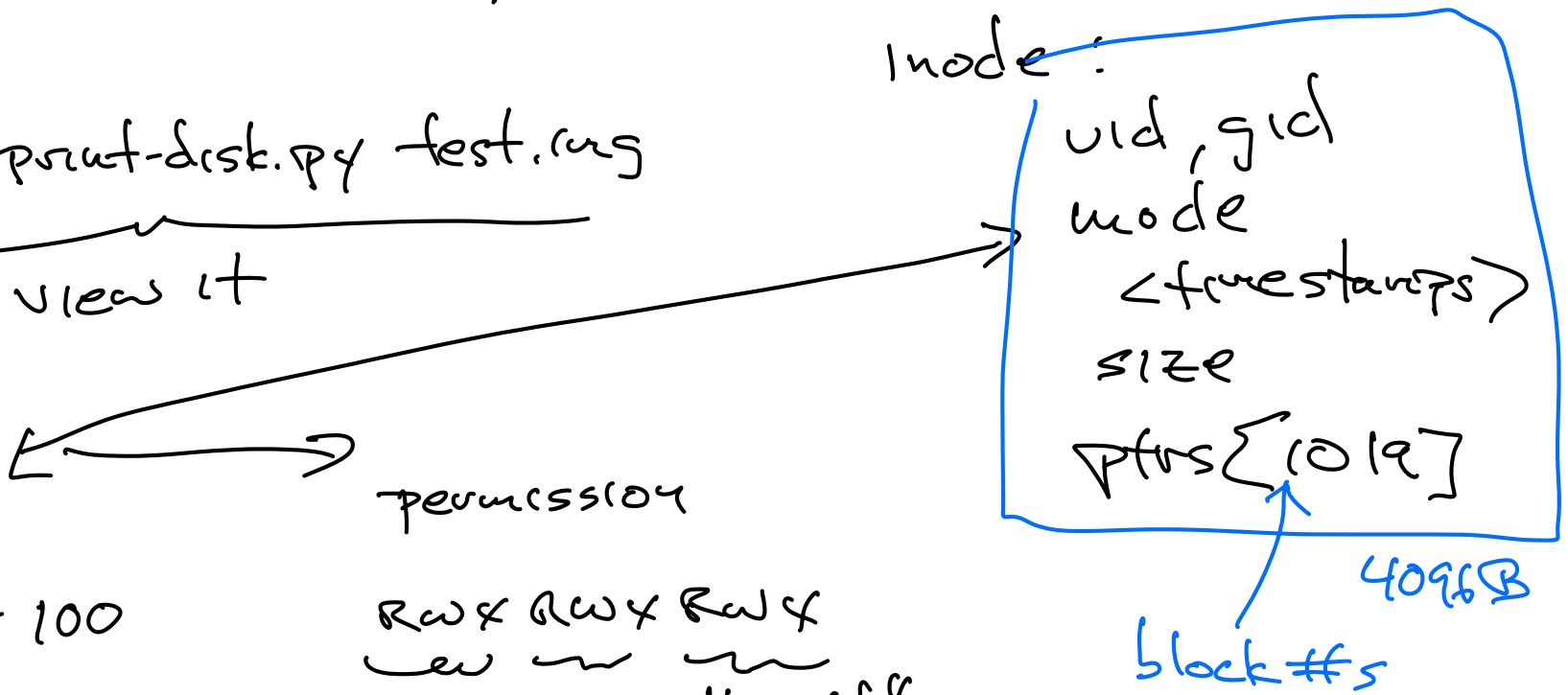
read-only
sort-of-unix
file system

```
python3 gen-disk1.py test.img
```

generate disk image

```
python3 print-disk.py test.img
```

view it



type

reg. file = 100

directory = 040

permission

RWX RWX RWX

user group "world"

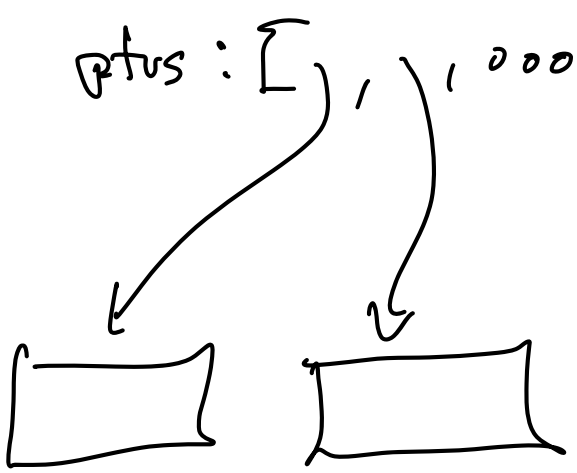
S_ISREG(mode) S_ISDIR(mode)

block_read(buf, block#, #blocks)

inode

size: 5000

ptrs: [, , 0000...]

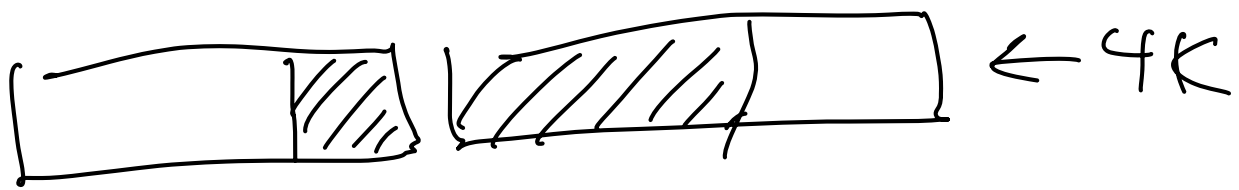


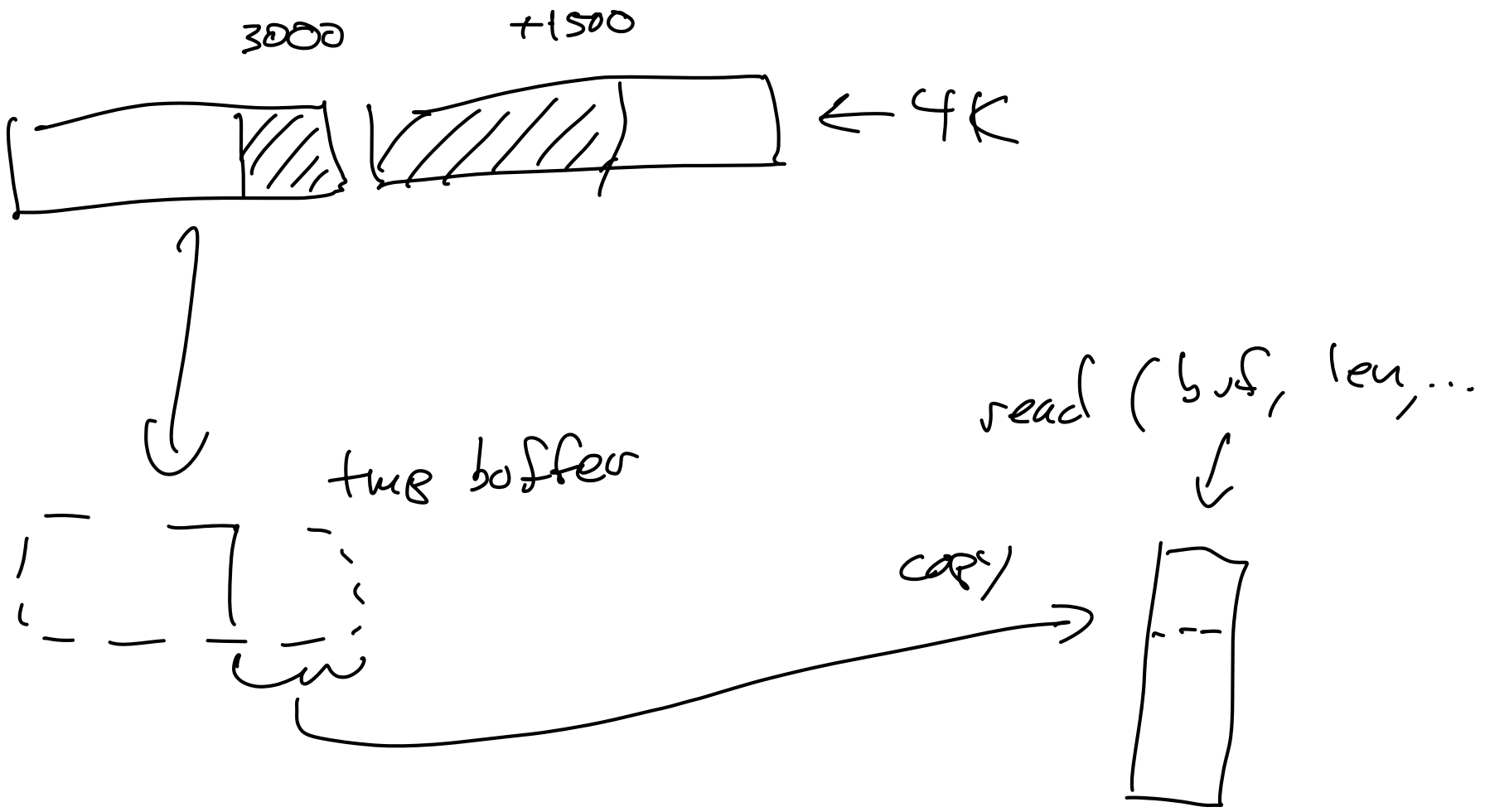
DIV_ROUND_UP(size, 4096)

$= 2$

$= \left\lfloor \frac{size + N - 1}{N} \right\rfloor$

read(offset, len) ← byte granularity




3000
+1500
← 4K
tmp buffer
read (buf, len, ...
copy