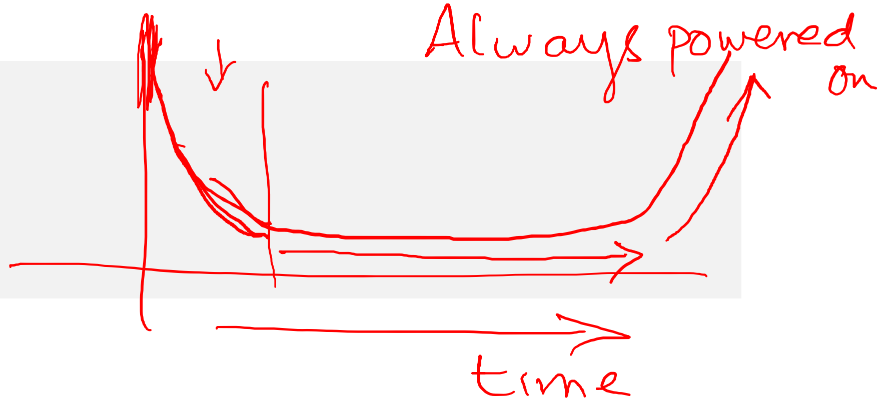
COL 362 & COL 632

Database Storage

21 Feb 2023

Usenix CMU

MTTF

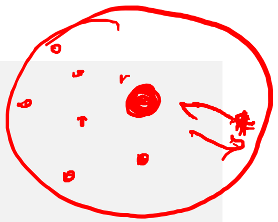


- MTTF = Mean Time To Failure
- If you had n disks, and you operated them for T hours before f of them failed then T*n/f
- *If you test a sample of 1,000 drives for a period of 1,000 hours and then one of them failed, then*

$$MTTF = \frac{1{,}000 \times 1{,}000}{1} = 10{,}00{,}000 \text{ hours} \sim 114 \text{ years}$$

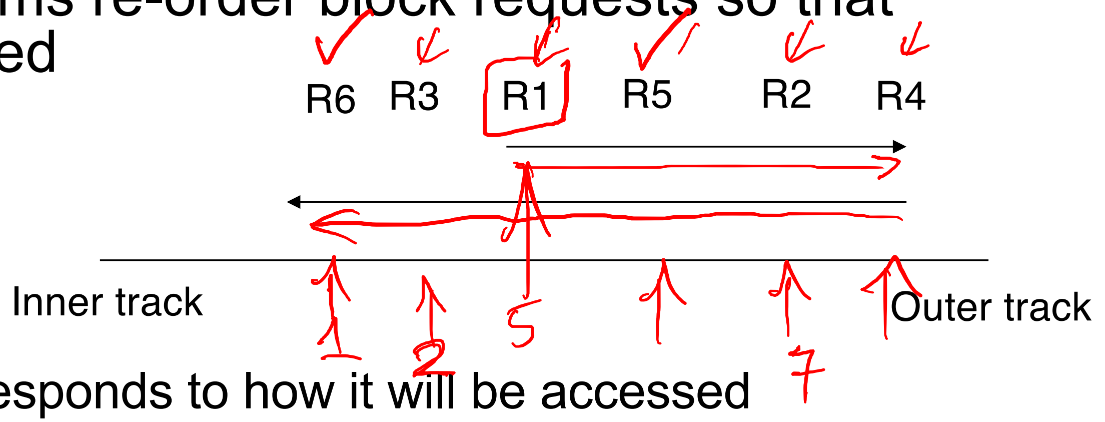
Now, if you had 114 drives, it is likely that at least one of them will fail after 1 year

Optimization of Disk-Block Access



Buffer manager

- Buffering
  - Fetching blocks from disk to satisfy future requests
  - DBMS implement their own buffering policies
- Read-ahead
  - Read extra blocks from a track in anticipation that they will be requested soon
- **Disk-arm-scheduling** algorithms re-order block requests so that disk arm movement is minimized
  - ***elevator algorithm***

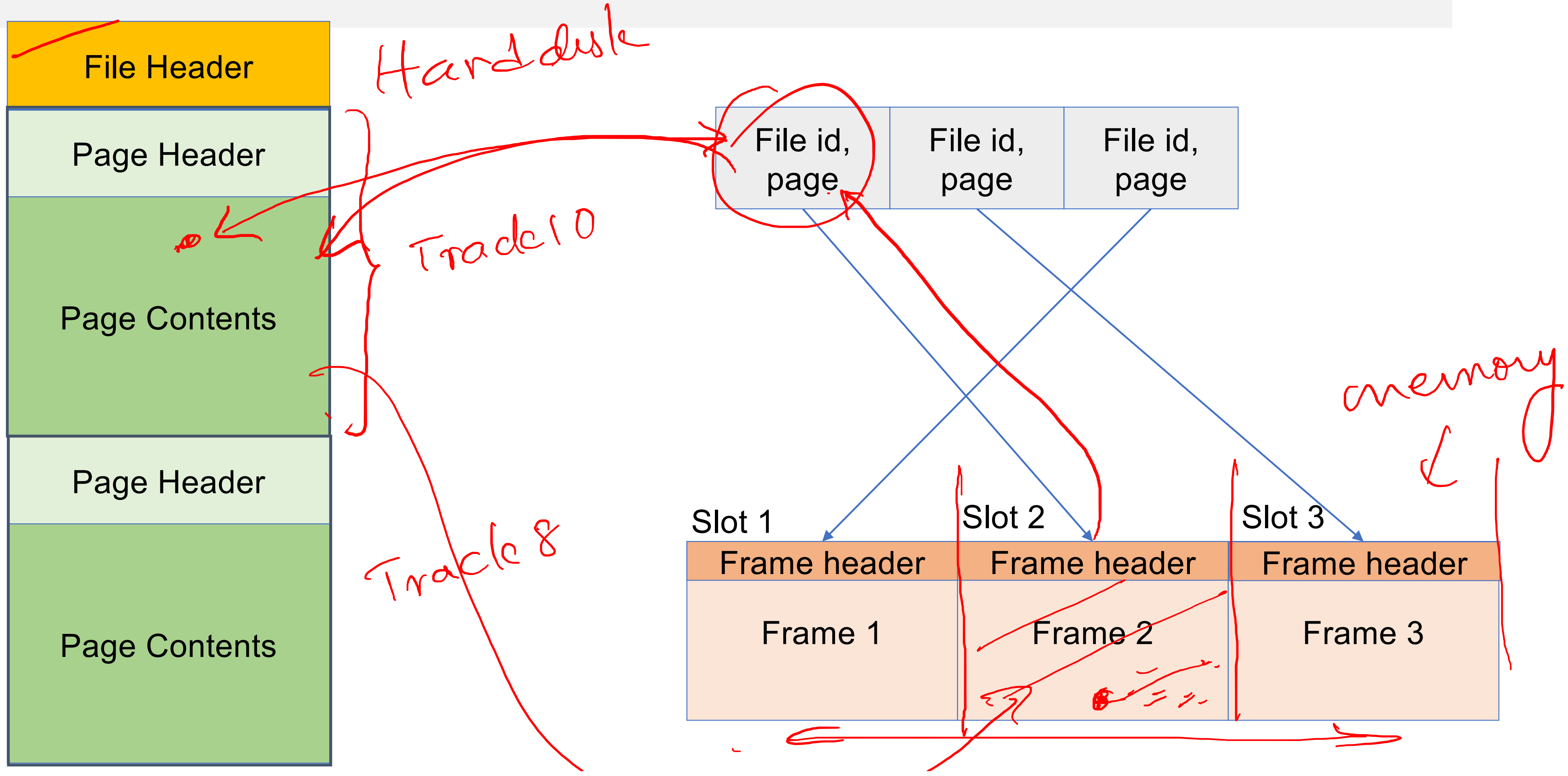


- File Organization
  - Organize files in a way that corresponds to how it will be accessed
  - Ex: contents of a file organized in consecutive blocks

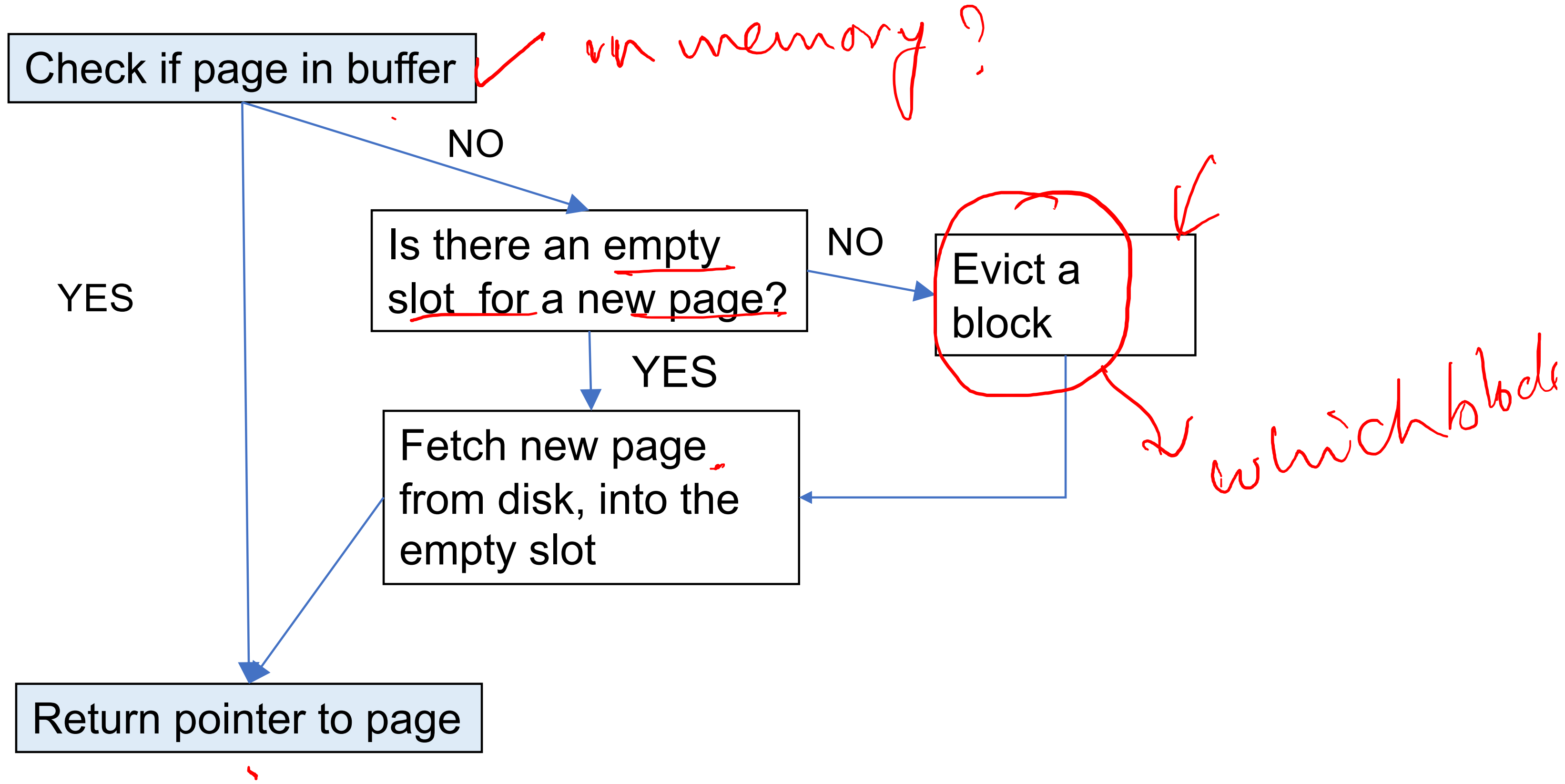
Storage Access

- Blocks are units of both storage allocation and data transfer.
  - Goal: minimize transfers from disk to memory
- Page is a unit of storage on disk that is structured
- Buffer – portion of main memory available to store copies of disk ~~blocks~~ pages
- Frame is a unit of storage in memory that is structured
- Buffer manager – subsystem responsible for allocating buffer space in main memory

Files and Buffers



Buffer Manager



Buffer Replacement Policies (1/2)

LRU-k

- Pinned blocks
  - Blocks which cannot be evicted until unpinned
- Forced output
  - Flush the block to disk
- Buffer replacement strategy
  - LRU, Clock, MRU, FIFO, etc.

Buffer Replacement Policies (2/2)

B1, B2, B3, B1, B4, B3, B2

	M1	M2	M3
B1	B1	-	-
B2	B1	B2	-
B3	B1	B2	B3
B1	B1	~~B2~~	B3
B4	B1	B4	B3
B3	~~B1~~	B4	B3
B2	B2	B4	B3

LRU – Least Recently Used

Nested-loop join

```
for each tuple i of instructor do
  for each tuple d of department do
    if i[dept_name] == d[dept_name] then
         <generate join tuple>
    end
  end
end
```

- For department blocks B1, B2, B3, B4 (and loop over for next block of instructor tuples)

LRU

	M1	M2	M3
B1	B1	-	-
B2	B1	B2	-
B3	B1	B2	B3
B4	B4	B2	B3

MRU – Most Recently Used

	M1	M2	M3
B1	B1	-	-
B2	B1	B2	-
B3	B1	B2	B3
B4	B1	B2	B4