

Storage

April 2, 2018

IO + Buffering

```
def Select(predicate, source)
  while(source.hasMoreTuples)
    in_buffer = source.fetch()
    while(in_buffer.hasMoreTuples)
      tuple = in_buffer.readTuple()
      if(predicate(tuple))
        out_buffer.output(tuple)
    if(out_buffer.isFull)
      out_buffer.flush()
```

IO + Buffering

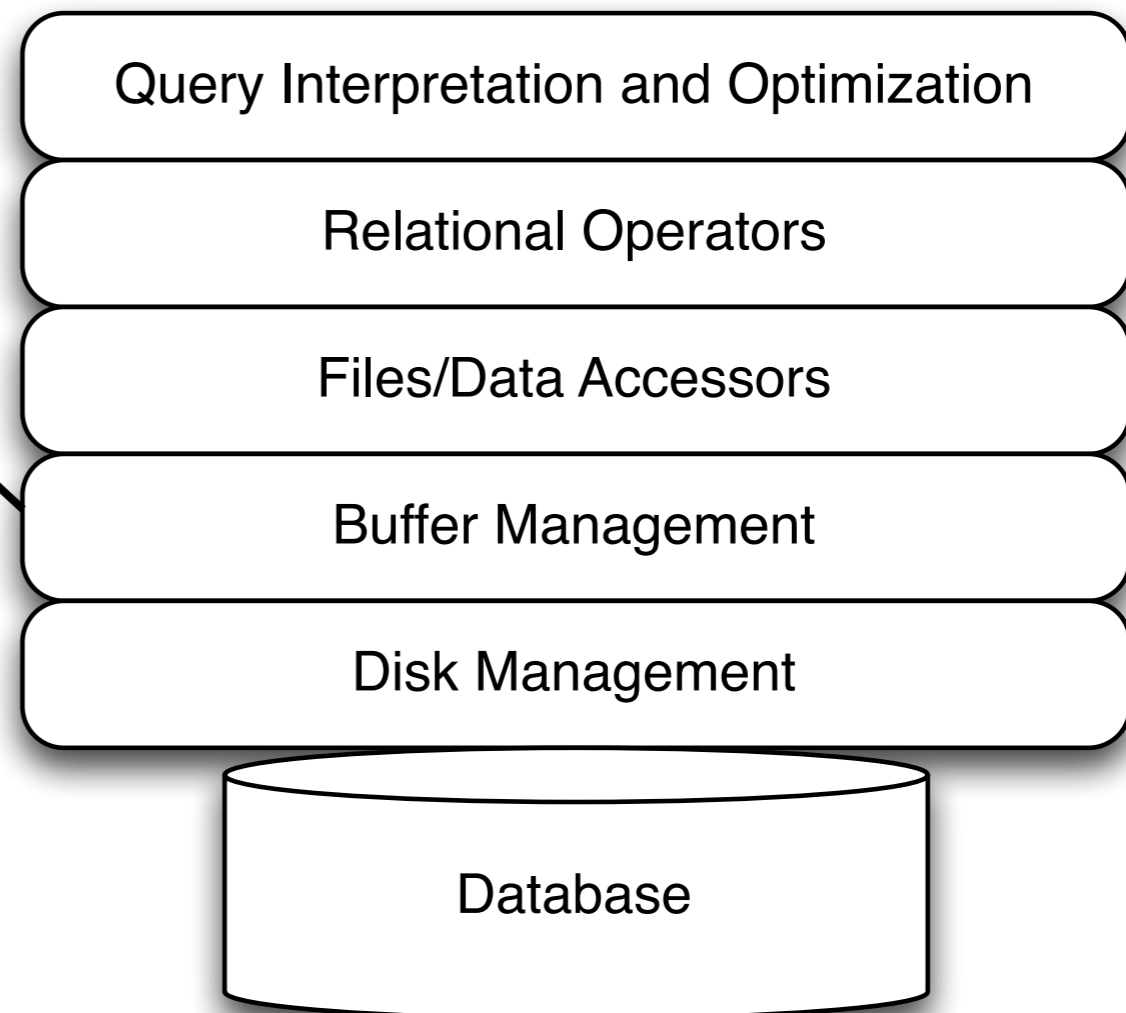
Generalize & Standardize!

Have a component that handles buffering!

The Buffer Manager

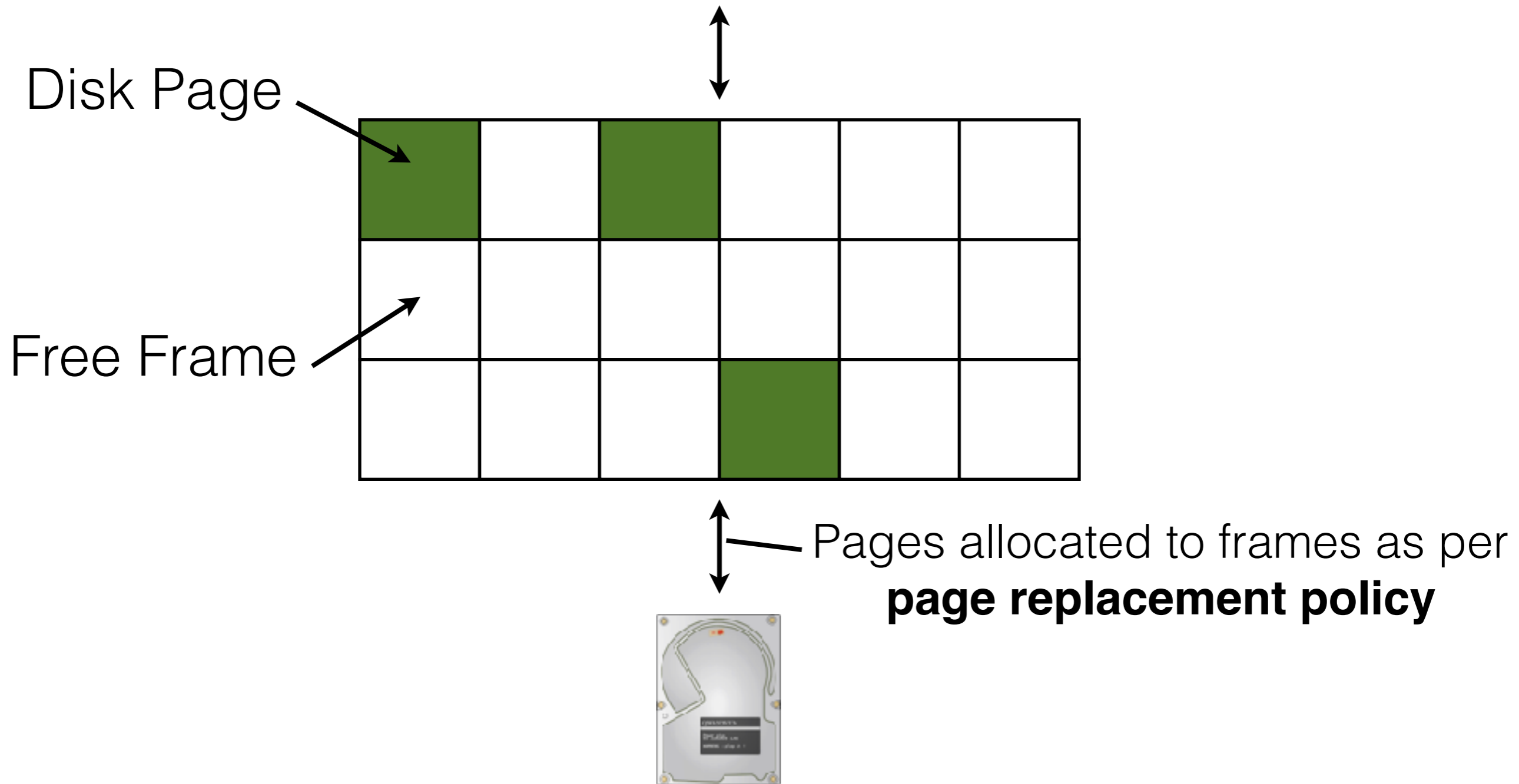
API

Allocate a page
Deallocate a page
Read from a page
Write to a page



The Buffer Manager

Higher levels of the DB



Pinned Pages

- Pinning a page indicates that it is being used.
- The requestor must unpin the page when done.
 - The requestor must also indicate whether the page has been modified (with a 'dirty' bit)
 - Dirty pages must be written to disk
- Pages may be requested multiple times
 - Use a pin count (reference count) to keep track.
- Concurrency Control/Recovery may require other operations when replacing a frame.

Buffer Replacement

- Frames are chosen for replacement by a **buffer replacement policy**.
- (e.g., LRU, MRU, Clock)
- Policy can have a big impact!
 - Depends on the access pattern.
- What is a worst-case scenario for LRU?

Buffer Replacement

- Frames are chosen for replacement by a **buffer replacement policy**.
- (e.g., LRU, MRU, Clock)
- Policy can have a big impact!
 - Depends on the access pattern.
- What is a worst-case scenario for LRU?
Hmmm... this sounds awfully familiar...

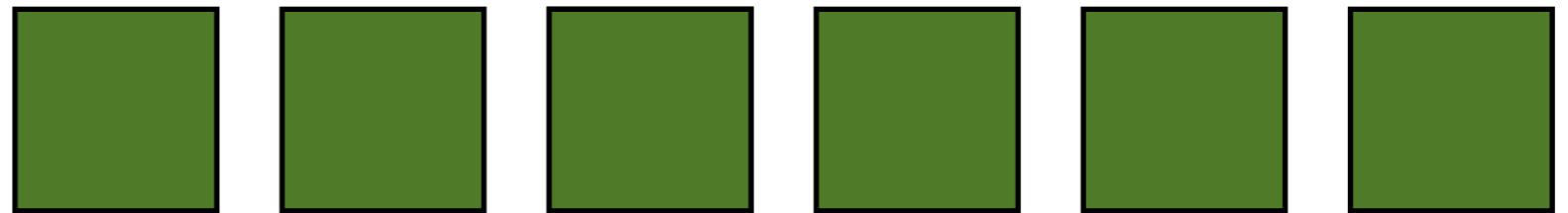
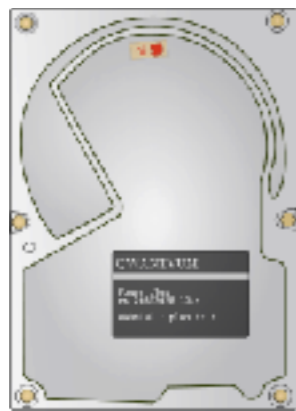
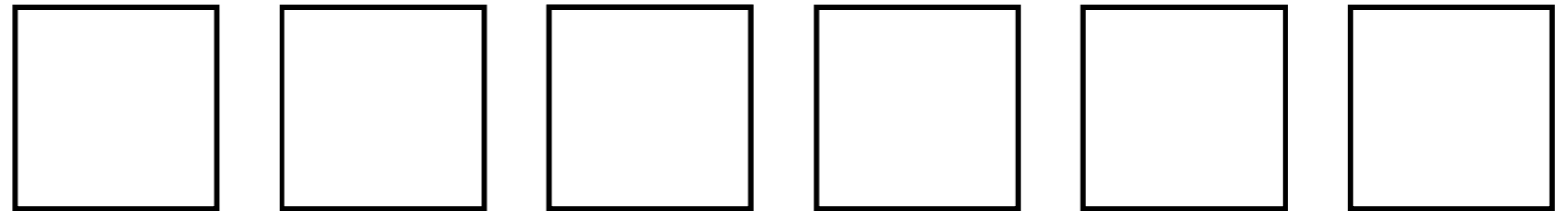
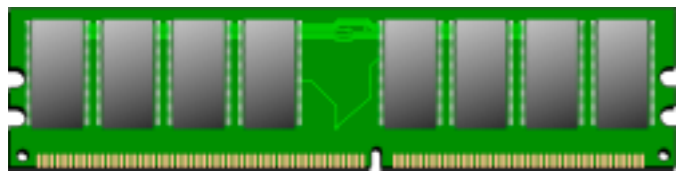
Hey... Oliver!

This sounds a lot like virtual memory!

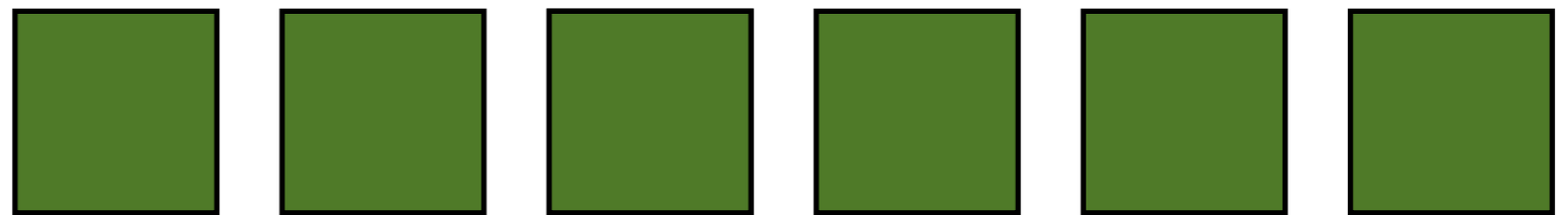
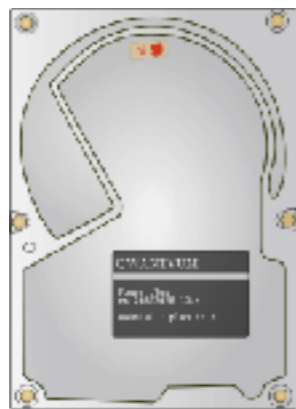
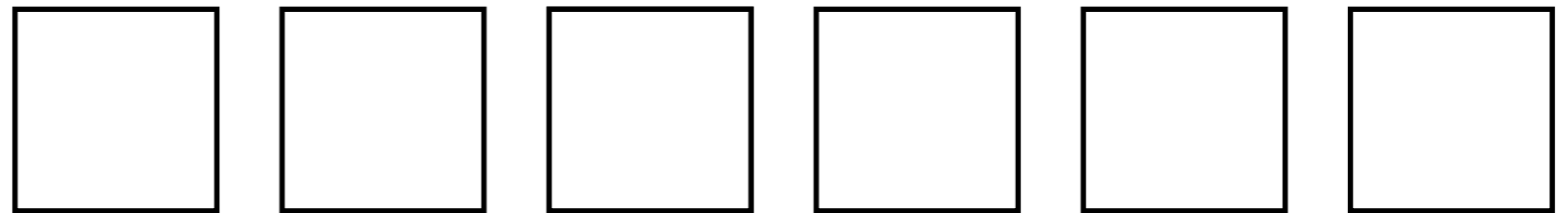
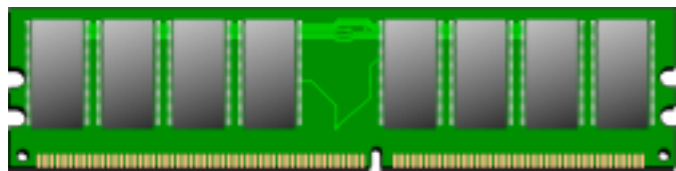
Buffer Managers vs Virtual Memory

- Not a huge difference
 - Many lightweight DBs use VMem as a buffer manager!
- Reasons to implement an explicit buffer manager:
 - Control when and how paging happens.
 - e.g., better/more efficient prefetching.
 - Control what gets paged in/out.
 - e.g., better knowledge of data access patterns.
 - Integrate additional memory layers (e.g., Network)

Example-OS Paging

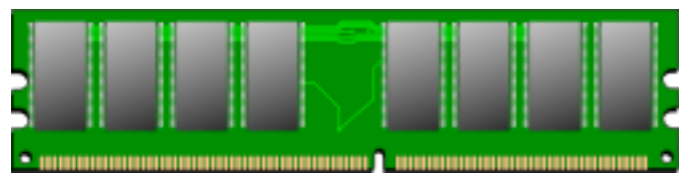


Example-OS Paging

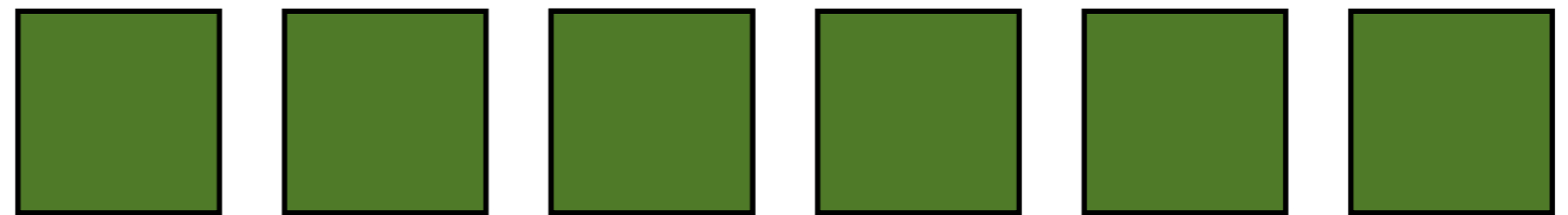
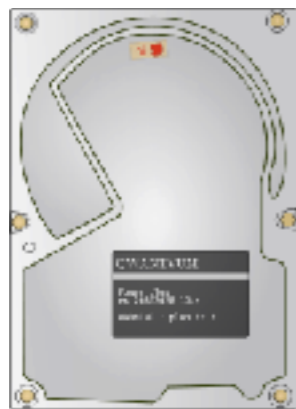
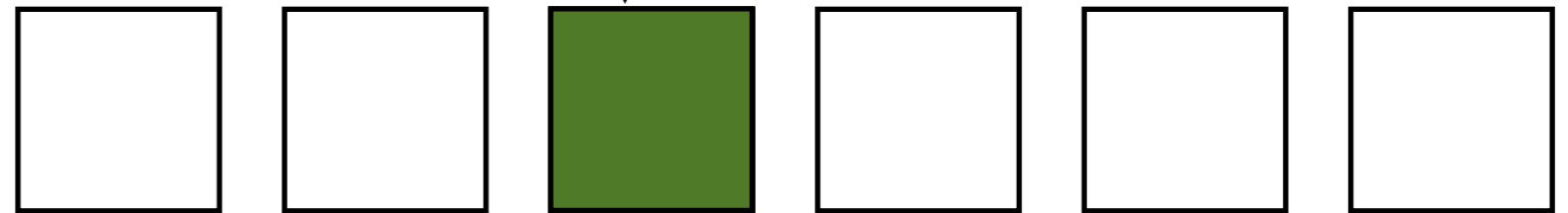


Index: 'Trees with ids in this range'

Example-OS Paging

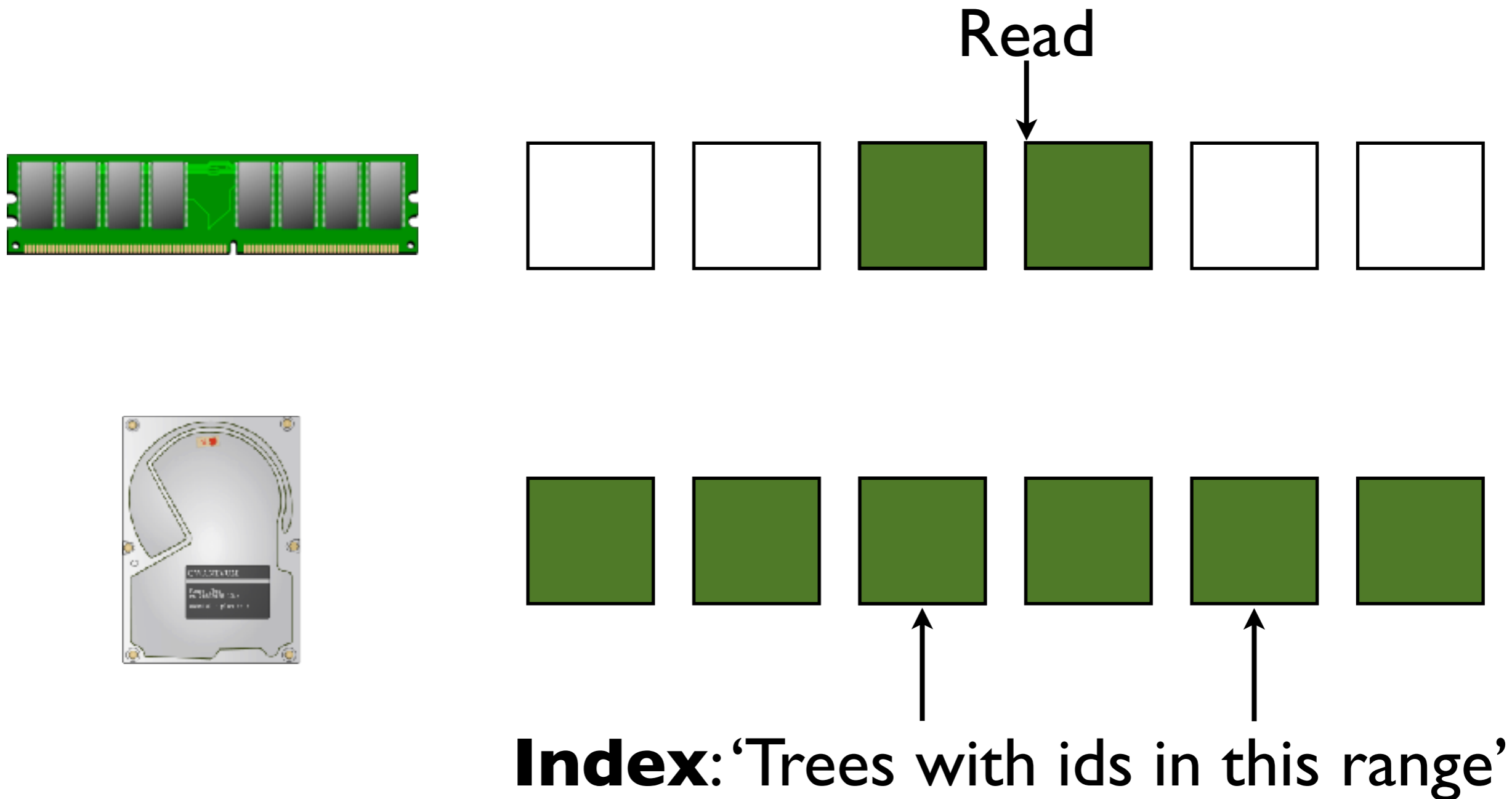


Read

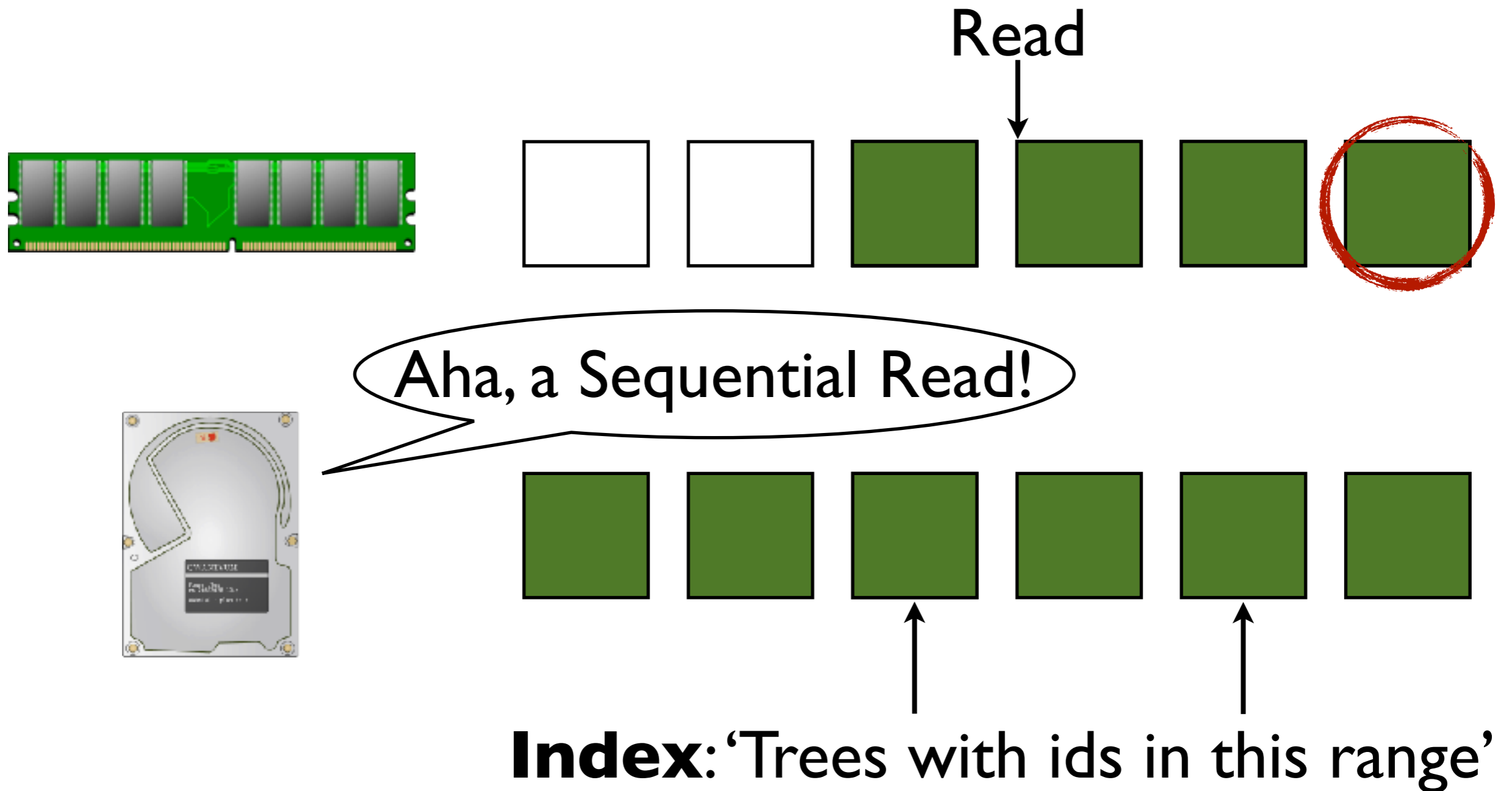


Index: 'Trees with ids in this range'

Example-OS Paging

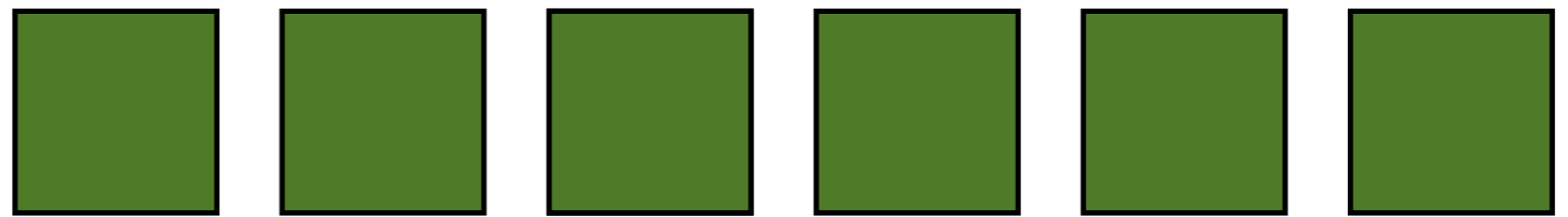
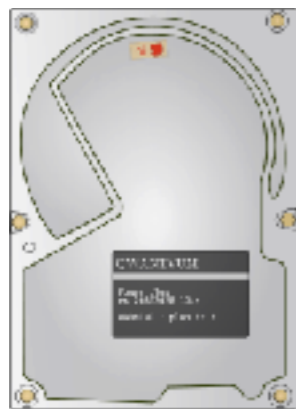
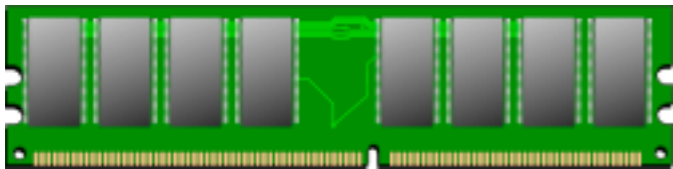


Example-OS Paging



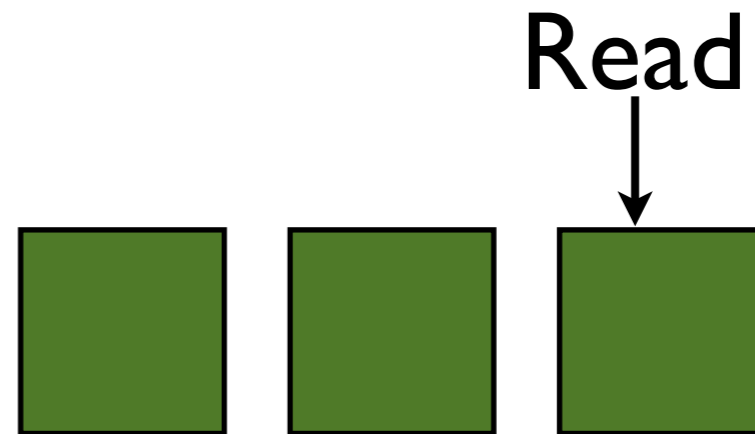
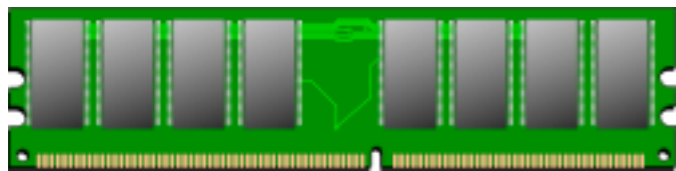
Example-DB Paging

Read
↓

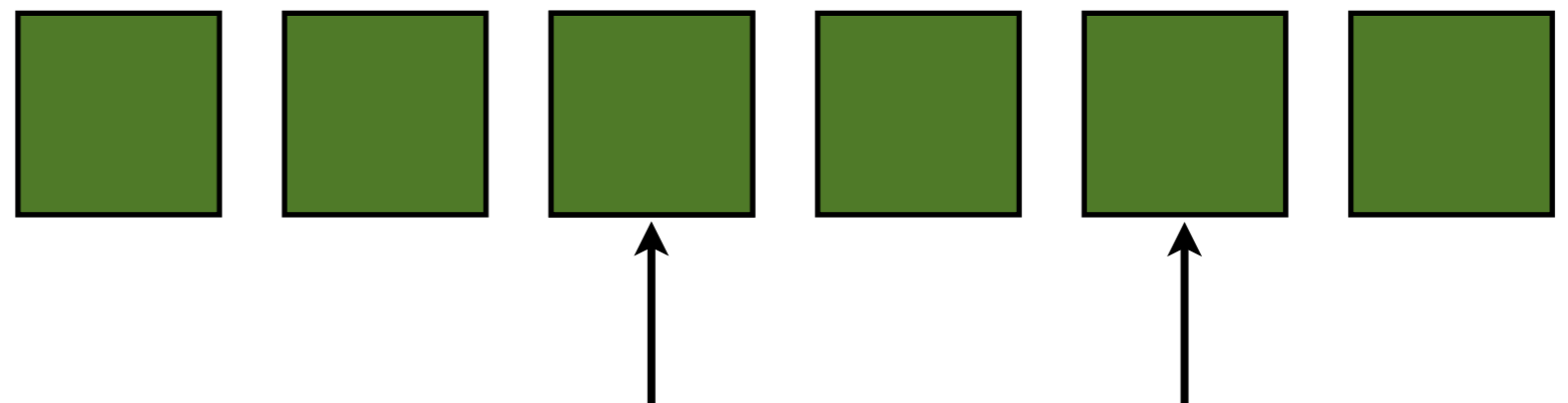
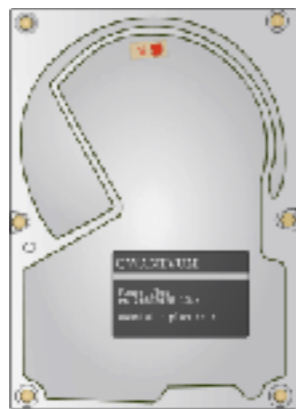


Index: 'Trees with ids in this range'

Example-DB Paging



Read in precisely what you need.



Index: 'Trees with ids in this range'