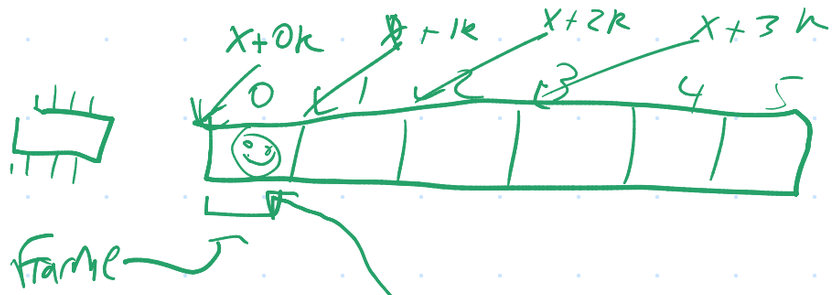
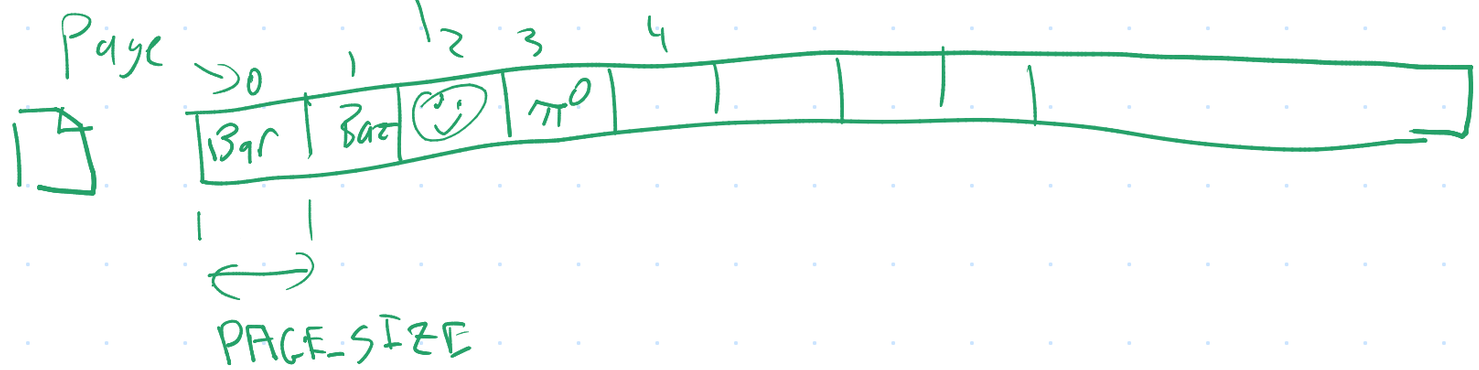
CSE 350

Advanced Data Structures

Lecture 3: The RAM Model is a Lie 🍰

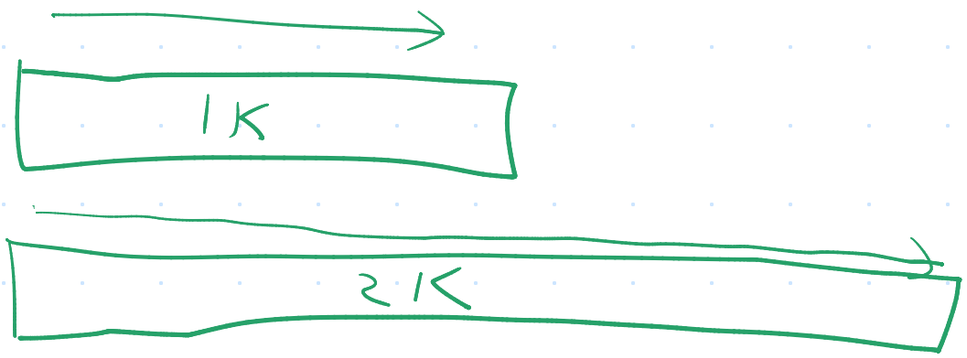
Page 2 @ X+0K



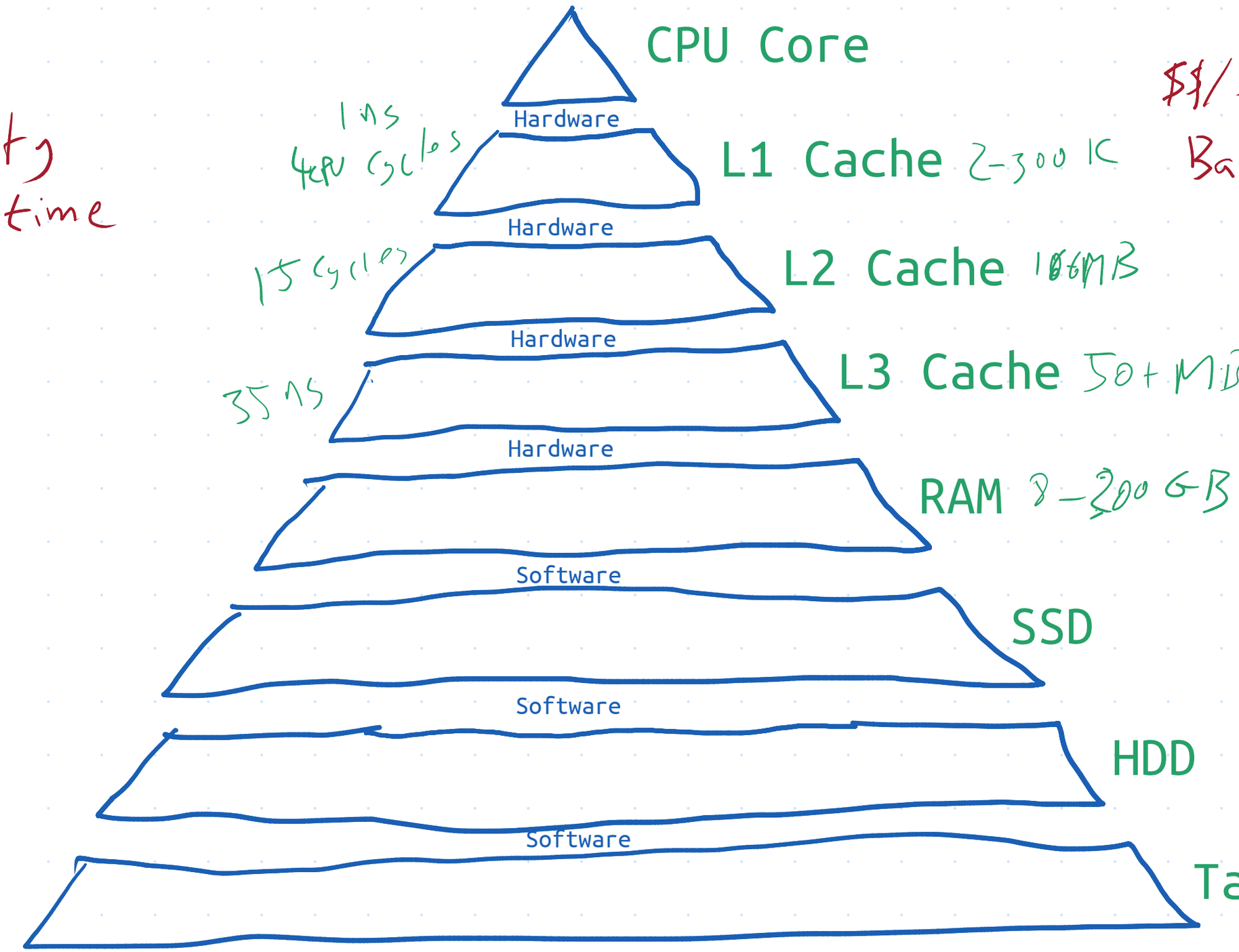


The "Ideal" Storage

- Unlimited Capacity
- Instant O(1) Access
- Unlimited Bandwidth
- Free
- Persistent



Capacity
access time



$/Byte
Bandwidth

Persistent

Access Scales

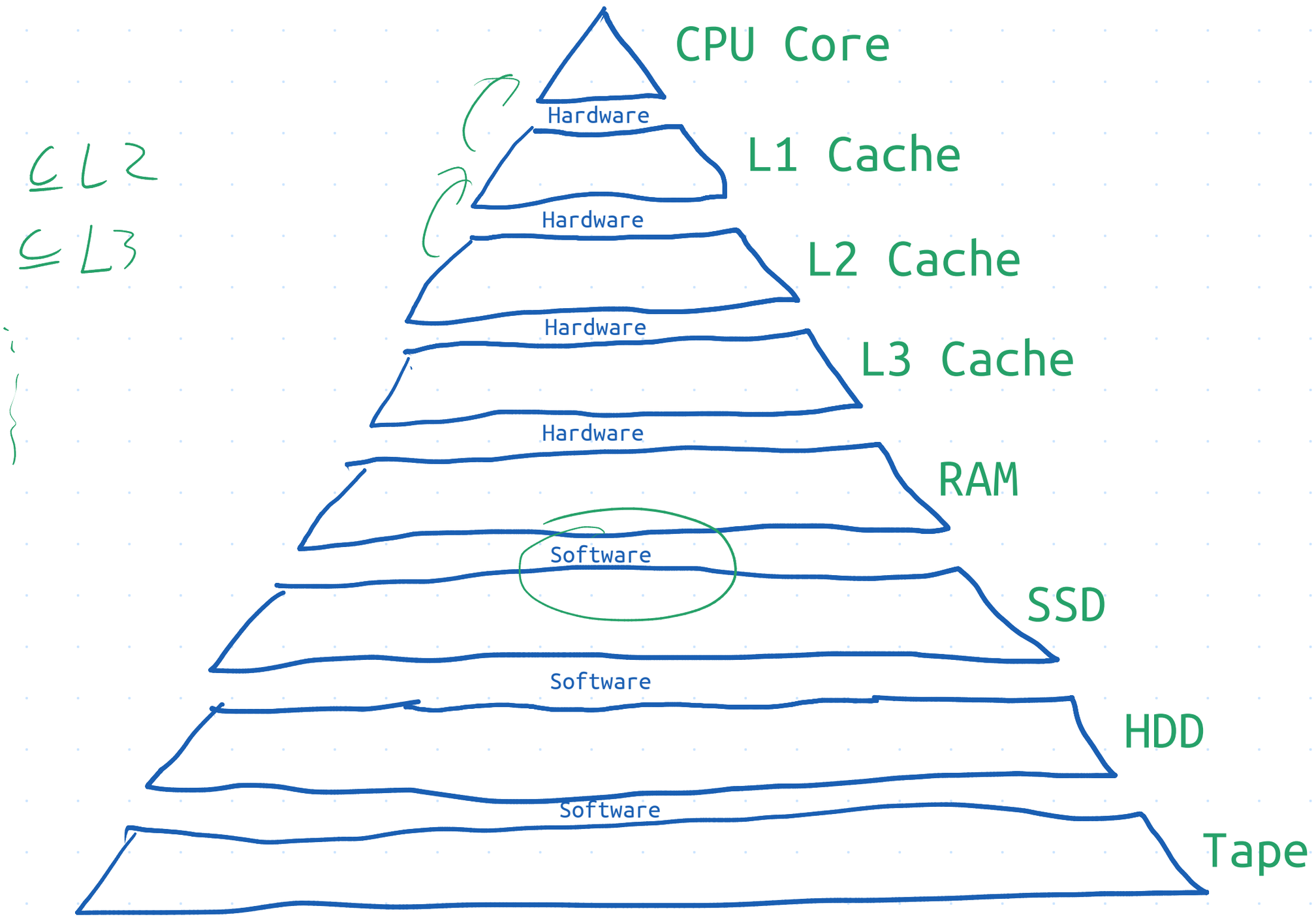
- CPU Register: Your phone ... is in your hand (2s)
- L1 Cache: ... is next to your bed (8s)
- L2 Cache: ... is on the other side of your room (30s)
- L3 Cache: ... is down the hall (70s)
- RAM: ... is on another floor/dorm (2 min)
- SSD: ... in Utica ... by foot (83 hours)
- HDD: ... in Portland ... by balloon
  ... the long way around (70 days)

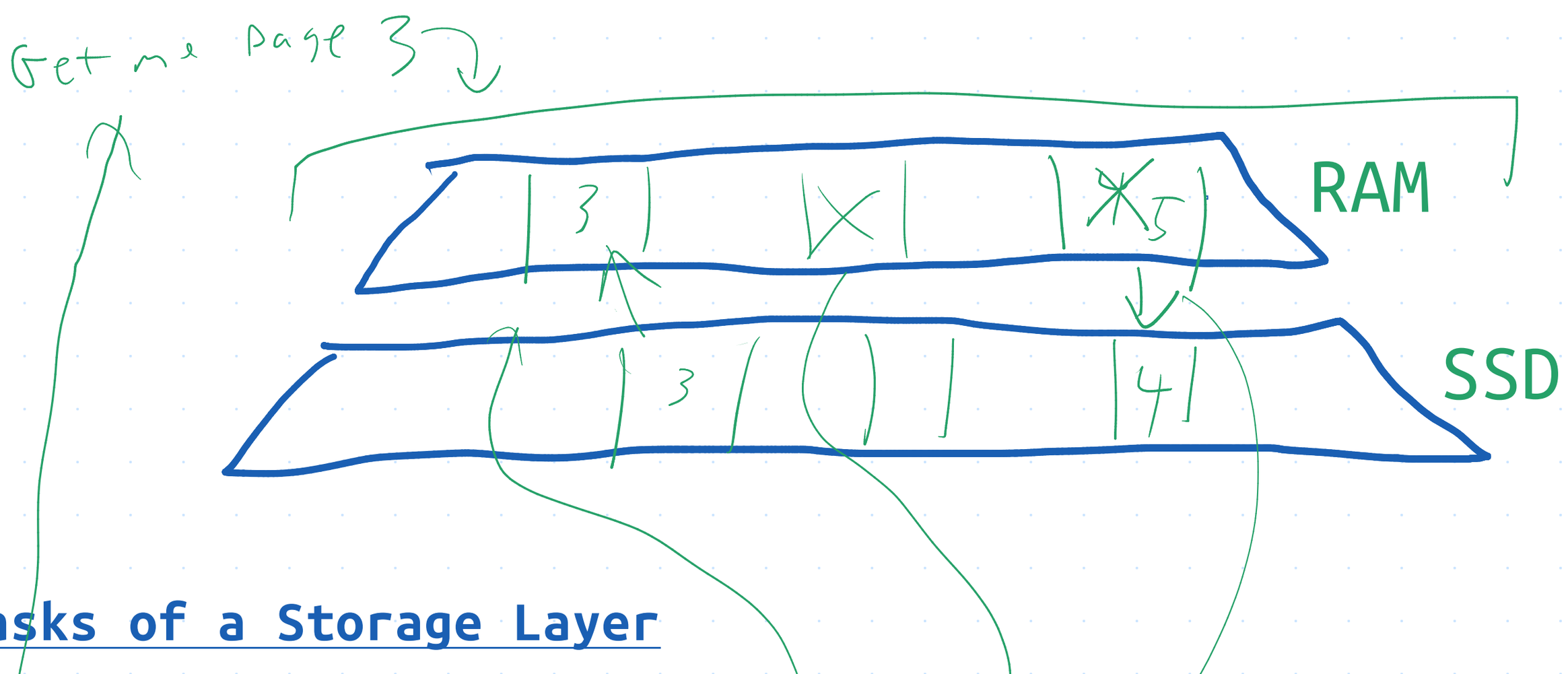
https://gist.github.com/jboner/2841832

The Inclusion Principle

$L1 \subseteq L2$

$L2 \subseteq L3$





Tasks of a Storage Layer

- Localizing Data Objects
- Caching Data from Lower Levels
- Data Replacement Strategies
- Writing Modified Data

The All Levels are Equal Principle

If:

- Technique X works on level Y

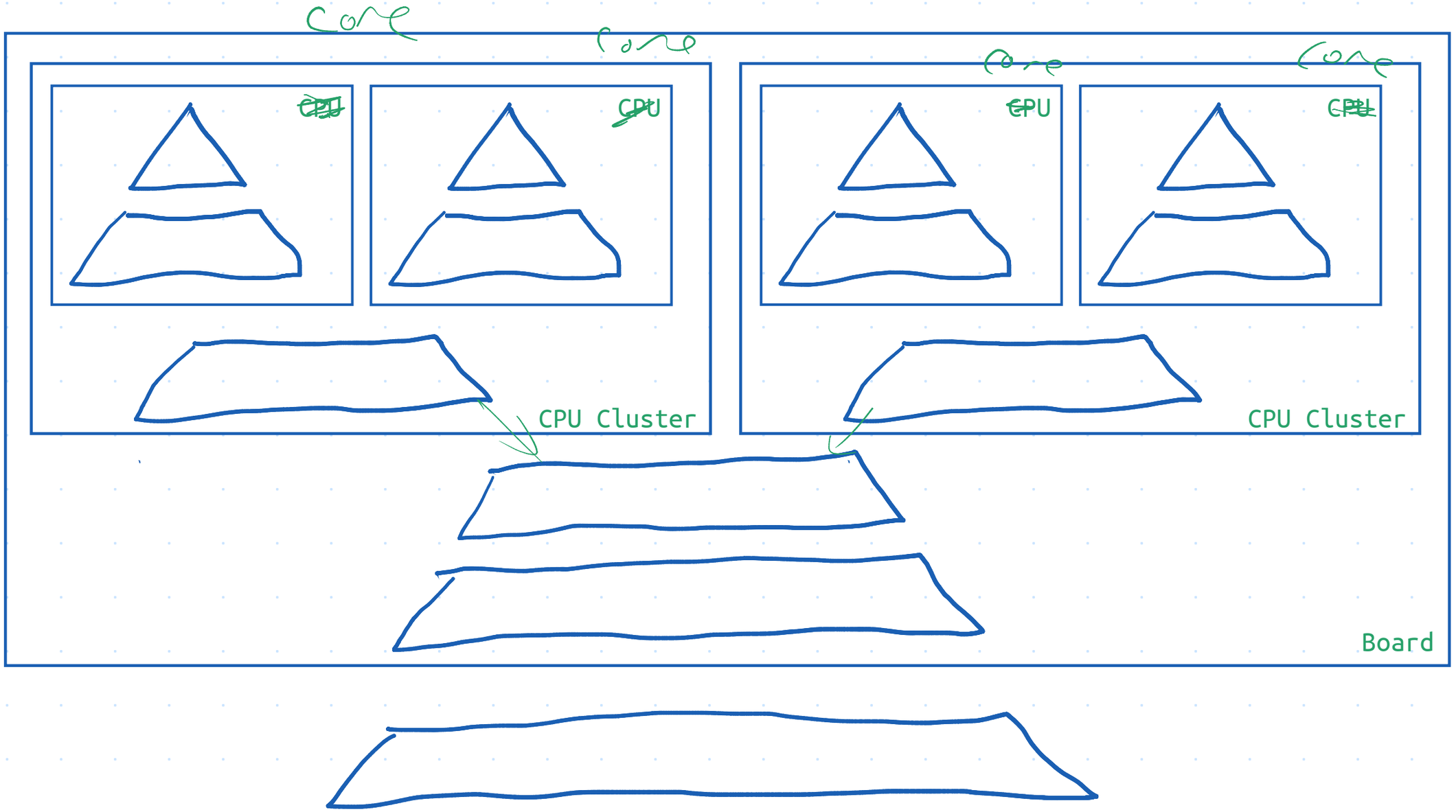
And:

- Technique X solves a problem for level Y

Then:

- Technique X will probably work on other levels

Multicore



CPU Core

L1 Cache

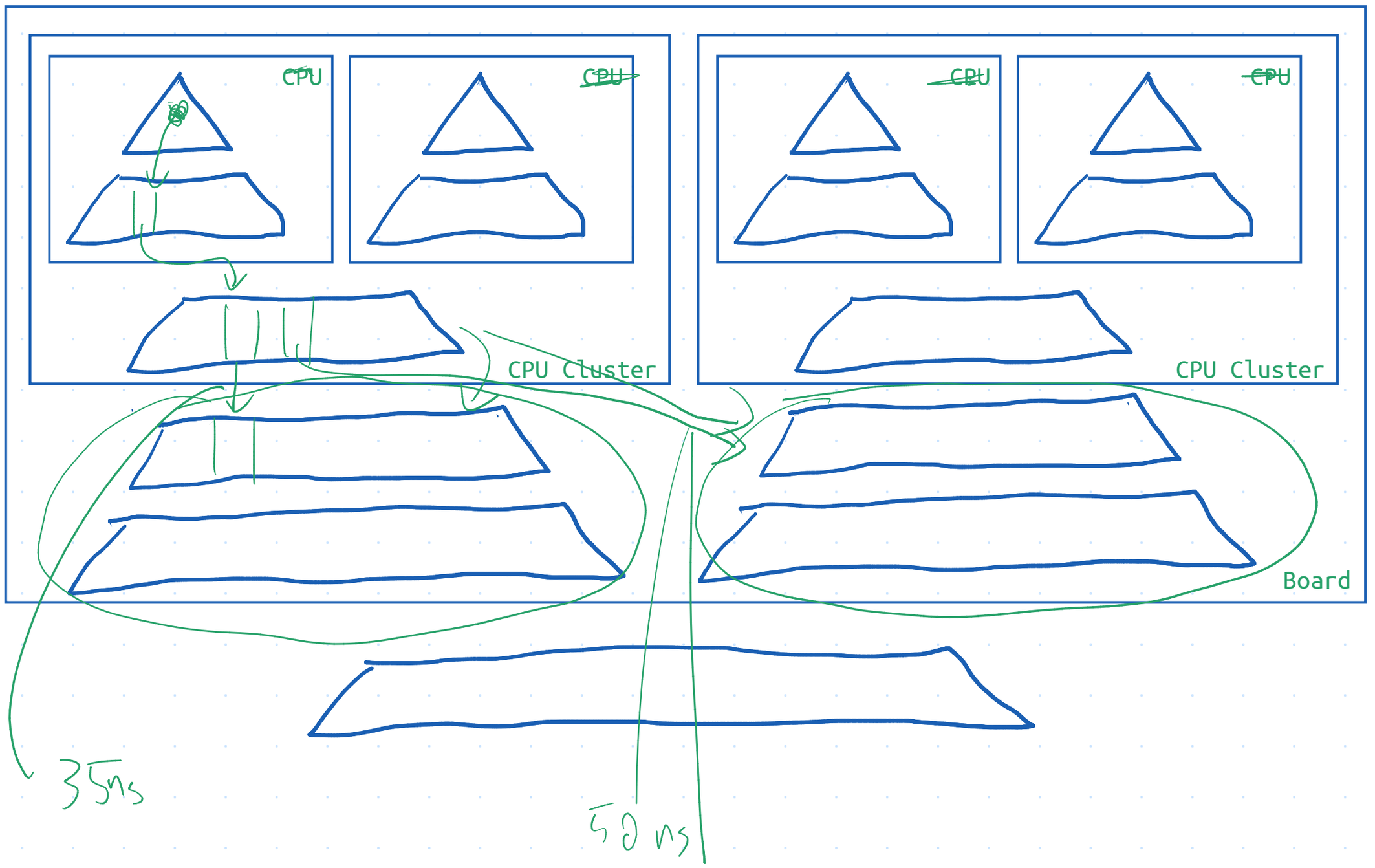
L2 Cache

L3 Cache

RAM

SSD

NUMA



CPU Core

L1 Cache

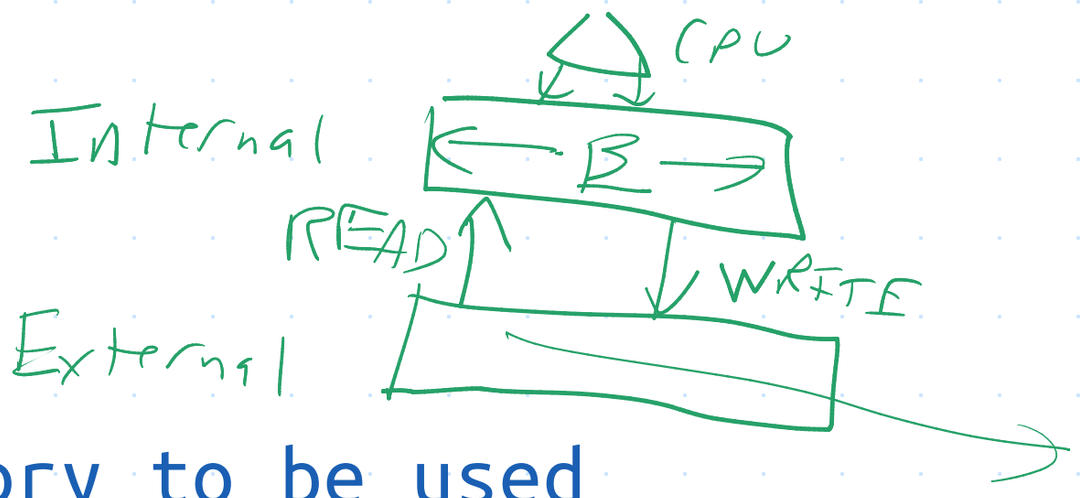
L2 Cache

L3 Cache

RAM

SSD

The EM Model



Internal

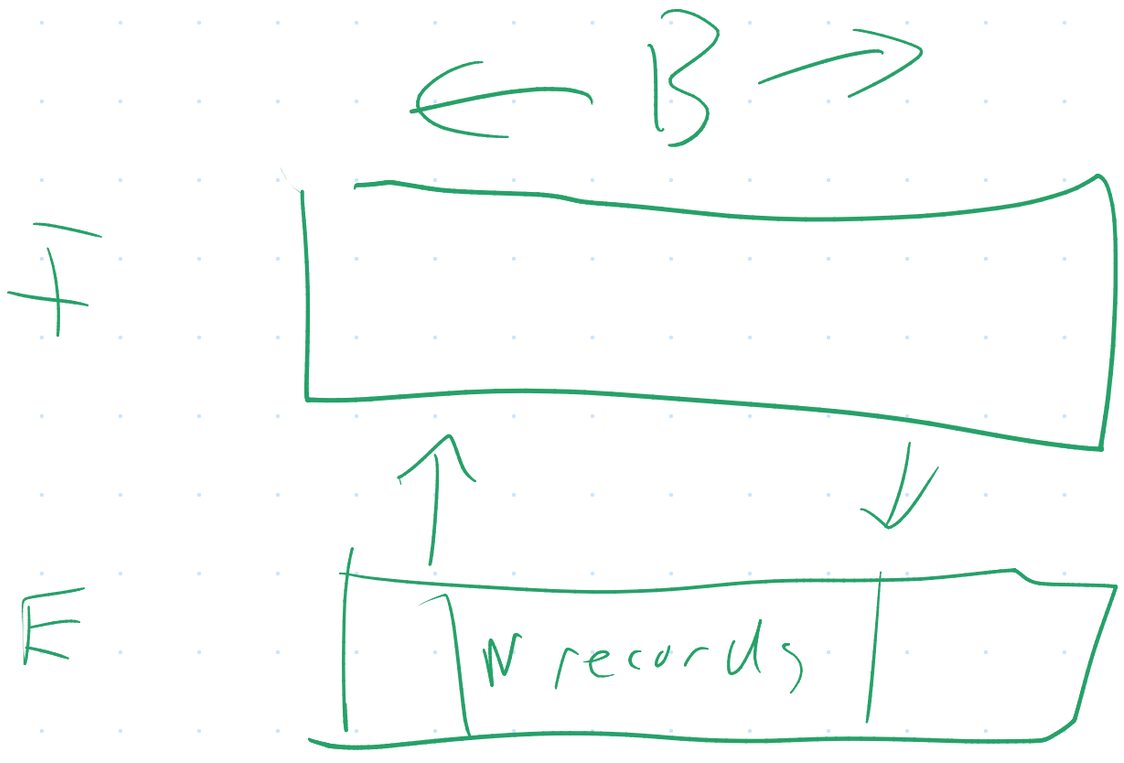
- Data must be in internal memory to be used
- The size of internal memory is limited
- Can WRITE a block of data to external memory

External

- The size of external memory is unlimited
- Can READ a block of data to internal memory
  - Old data in that part of internal memory is lost

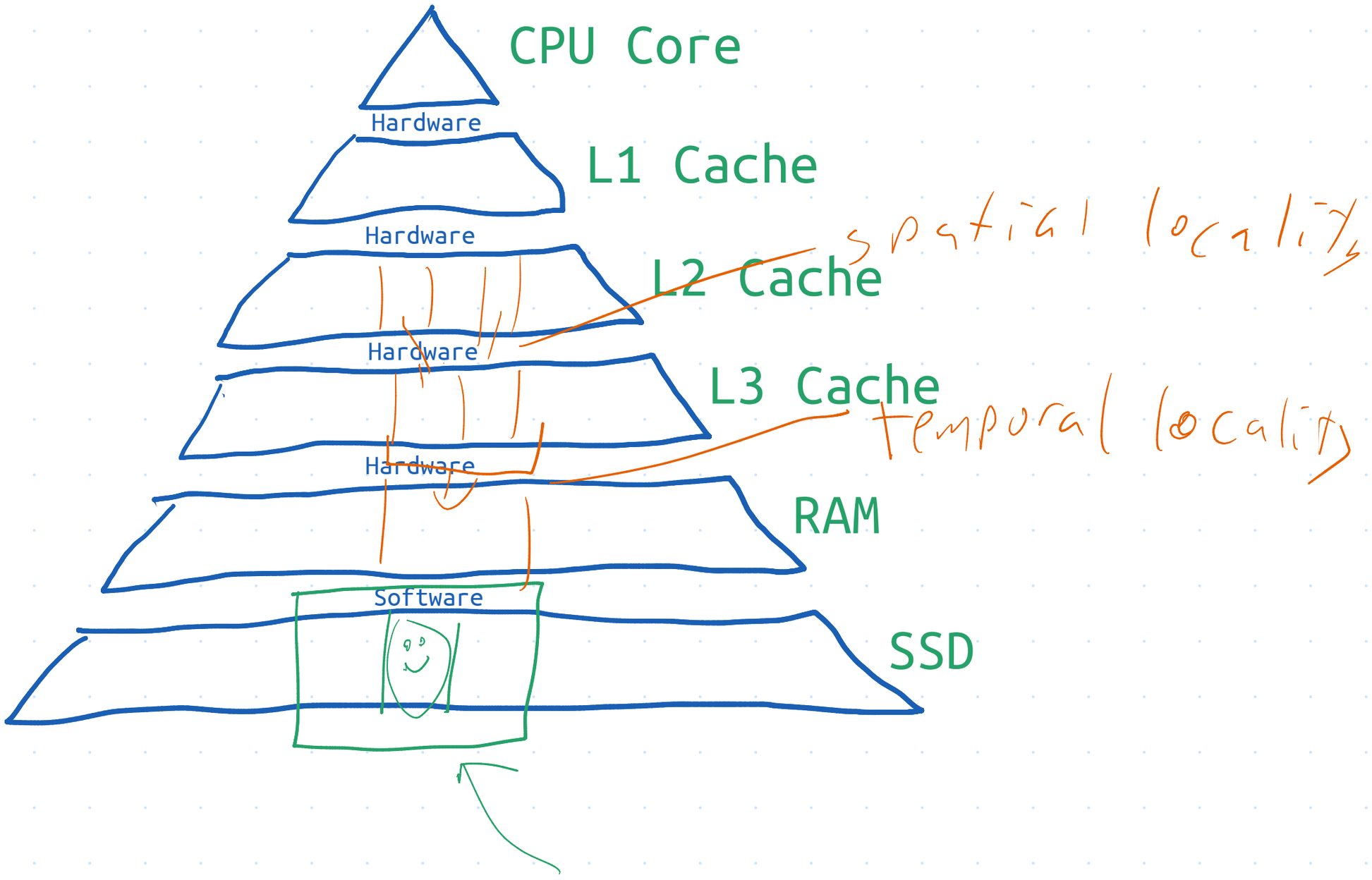
Algorithm Measures

- Number of steps
- Number of calls to READ+WRITE (IOs)
- Minimum size of internal memory (Working Set Size)

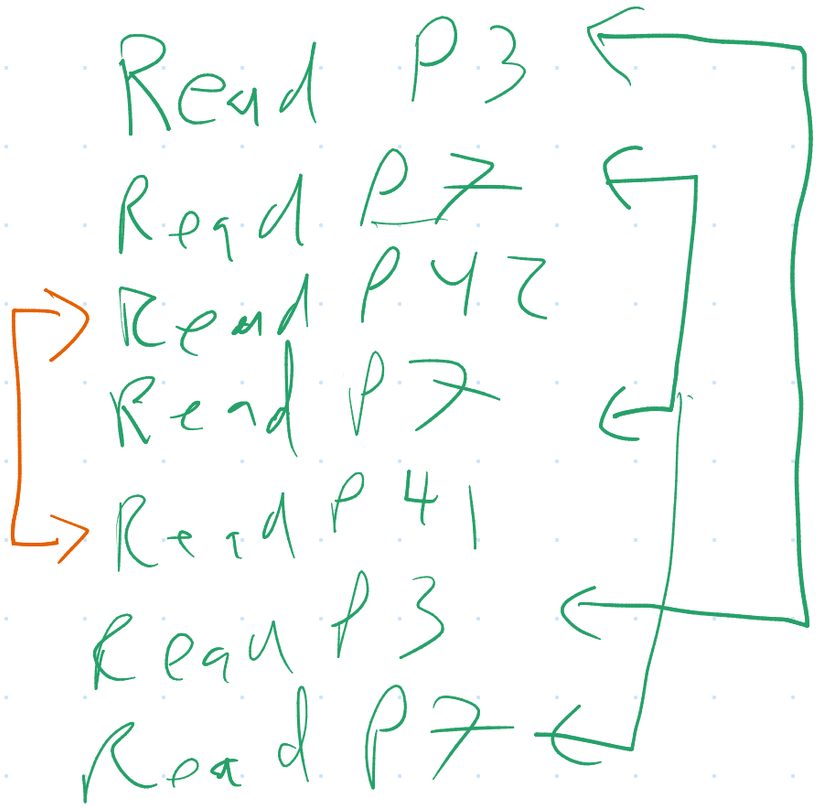


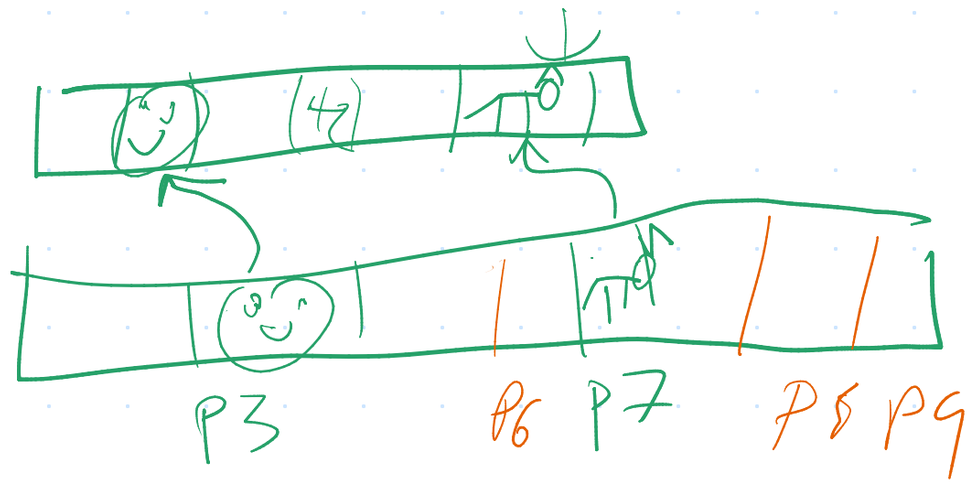
$B = O(N)$

	IO	Mem
Quicksort		
Read	$O(N)$	$O(N)$
Apply Algo	$O(0)$	$O(N)$
Write	$O(N)$	$O(N)$


CPU Core
Hardware
L1 Cache
Hardware
L2 Cache
spatial locality
Hardware
L3 Cache
temporal locality
Hardware
RAM
Software
SSD

Temporal Locality

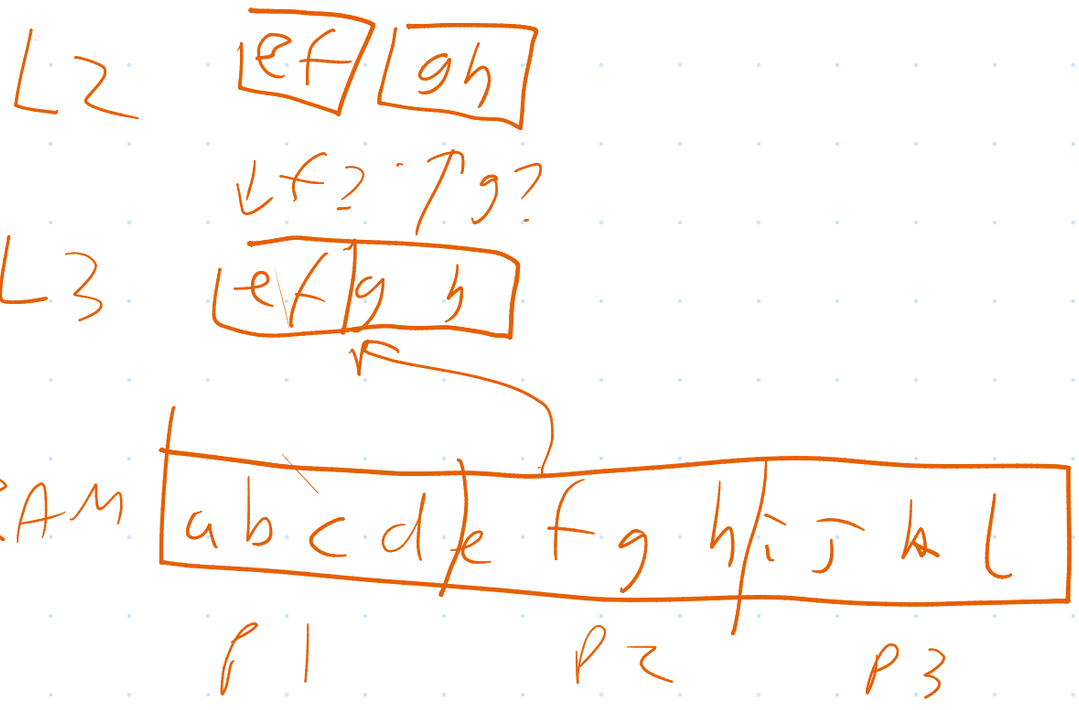




Spatial locality →

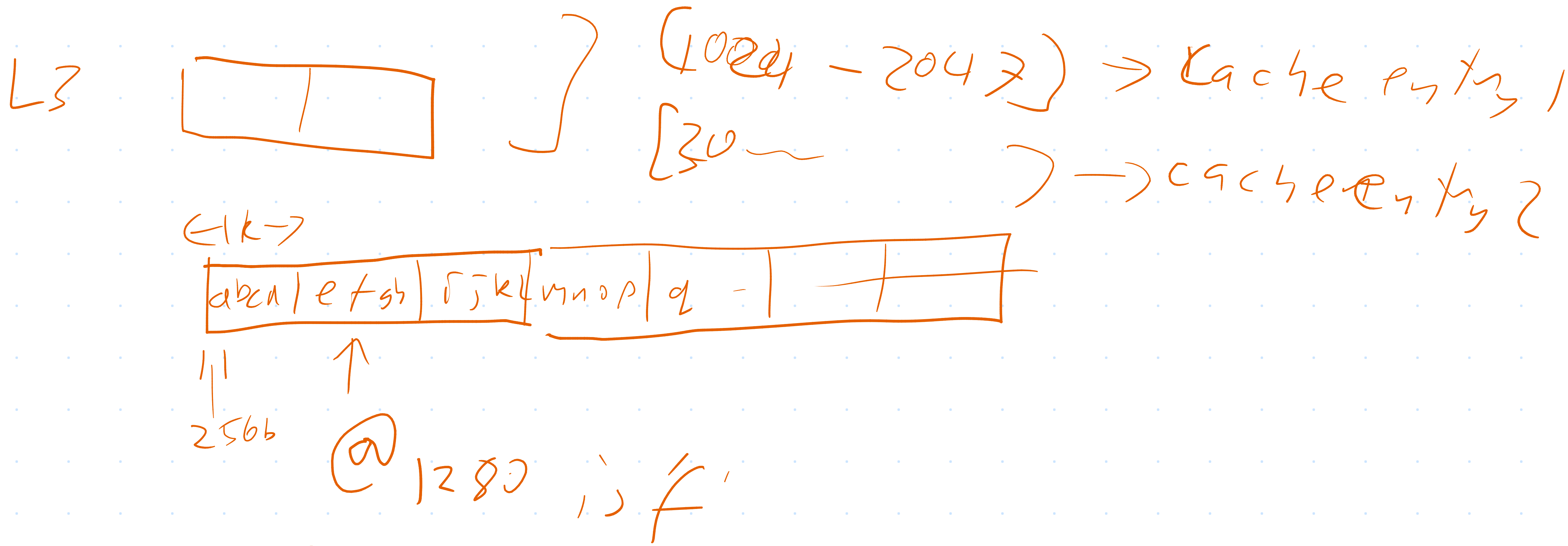
$$d(p_i, p_j) = |i - j| \leq \frac{S}{X}$$

f? g



1P = 2 records

1P = 4 records


L3
[1024 - 2047] → cache entry 1
[30
→ cache entry 2
←1k→
abcd
efgh
ijkl
mnop
q
256b
@1280