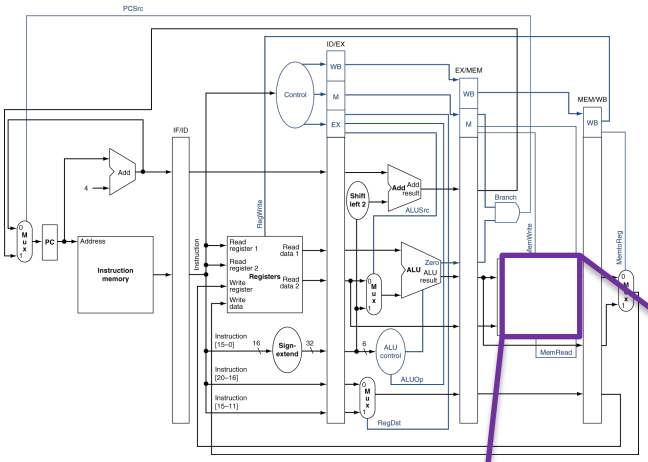
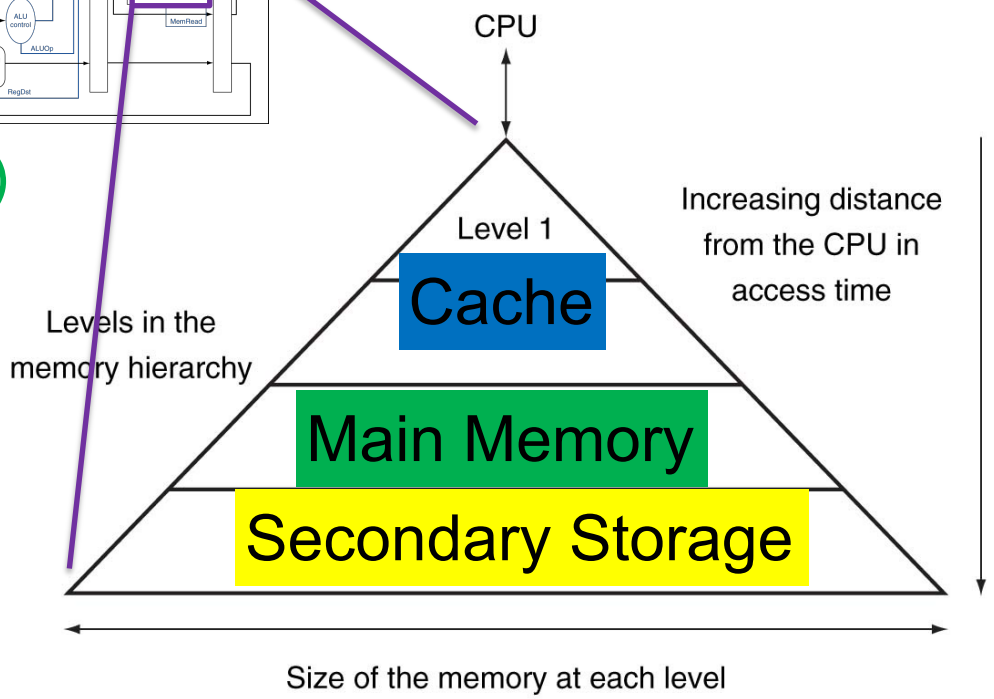
Virtual Memory

Memory Hierarchy

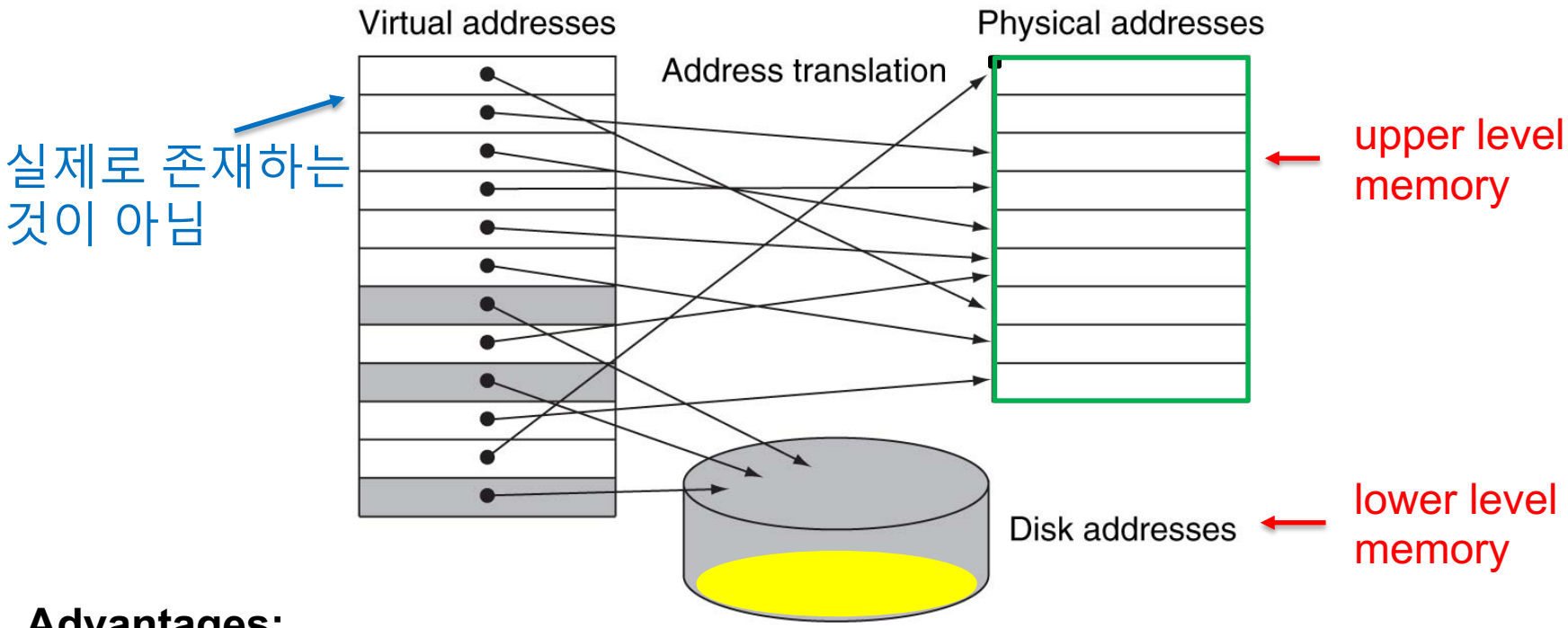


lw $4, 0($3)

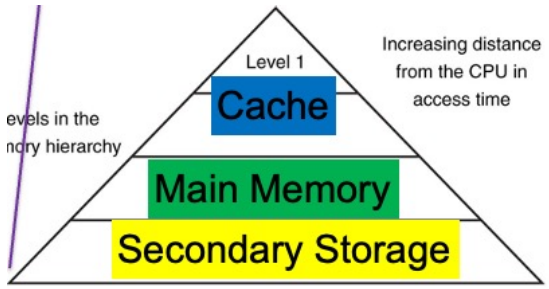


Virtual Memory

- **Main memory can act as a cache for the secondary storage (disk)**



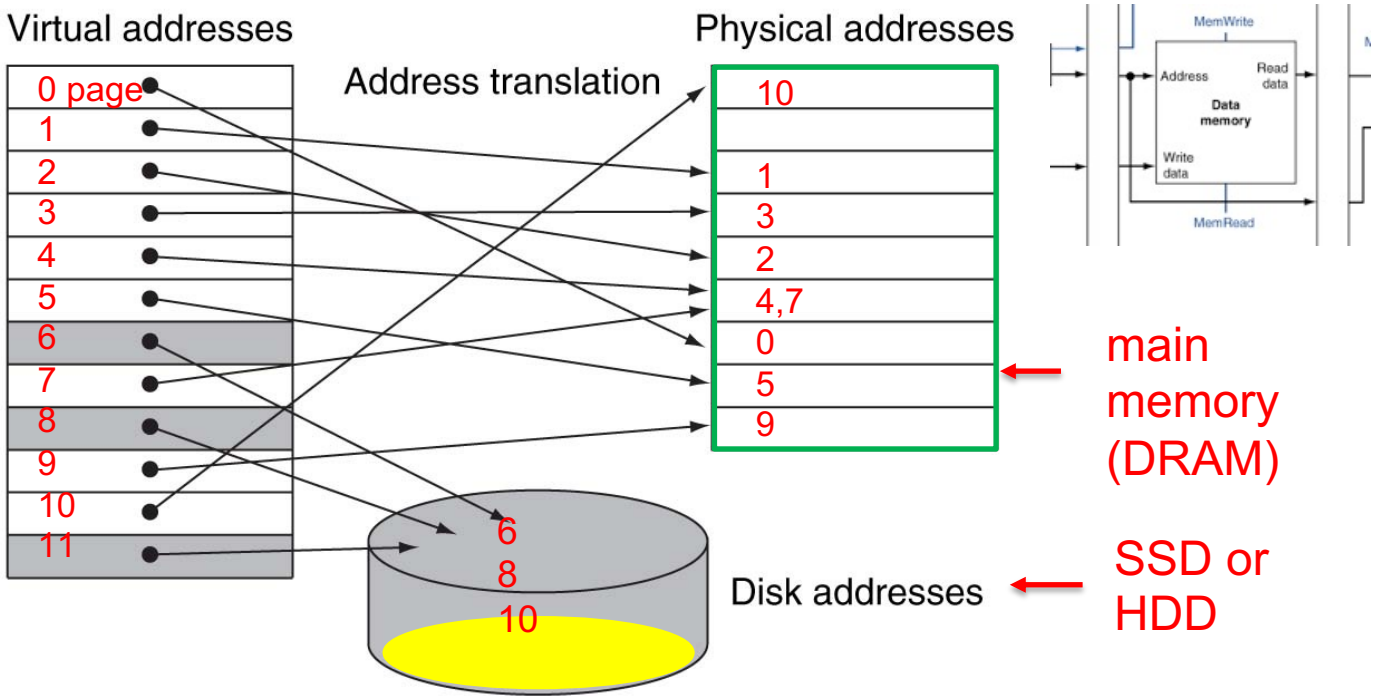
- **Advantages:**
  - **illusion of having more physical memory**
  - **program relocation**
  - **protection**



Virtual Memory

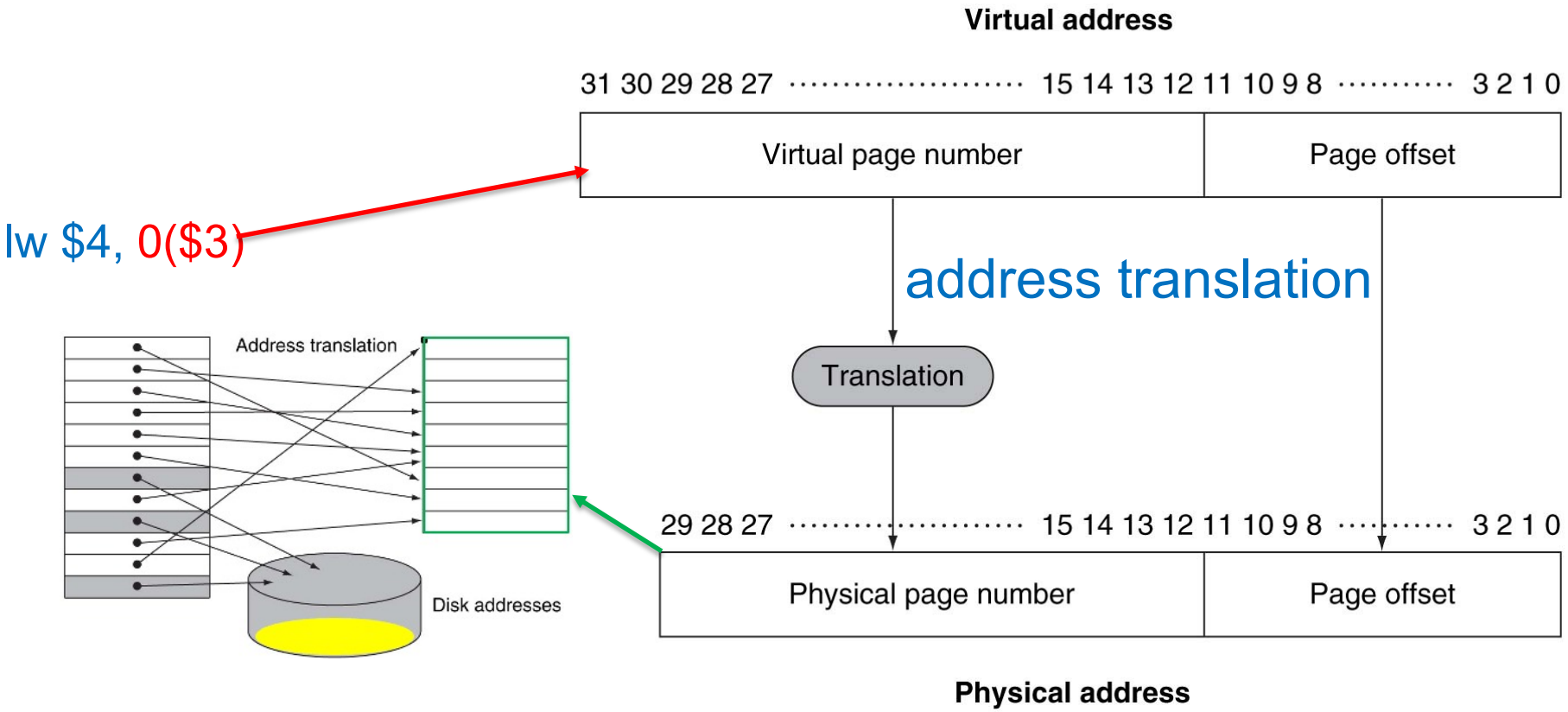
program uses virtual address

lw $4, 0($3)



Pages: virtual memory blocks

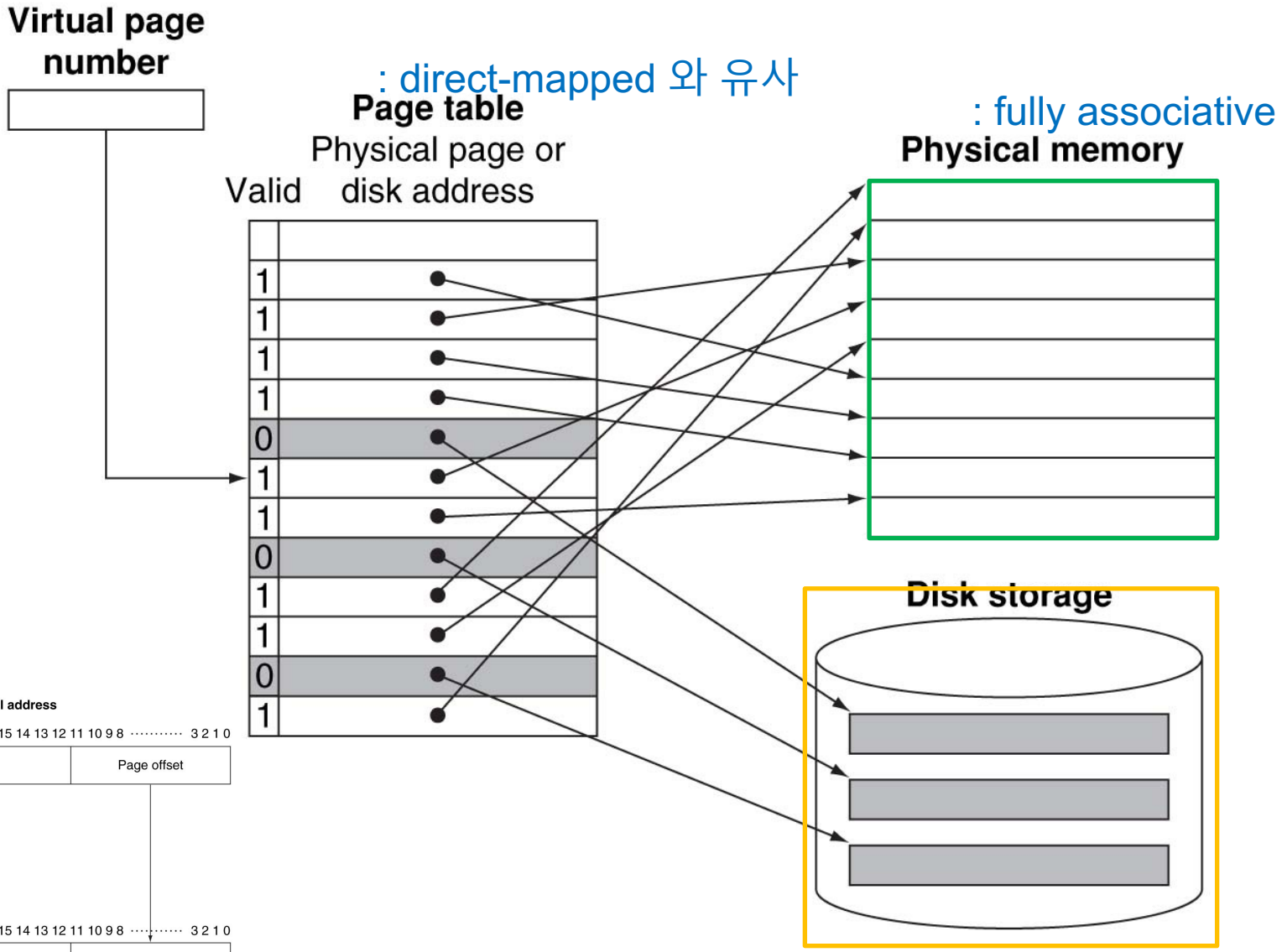
- **Page faults: the data is not in memory, retrieve it from disk**
  - **huge miss penalty, thus pages should be fairly large (e.g., 4~16KB)**
  - **reducing page faults is important (LRU is worth the price)**
  - **can handle the faults in software instead of hardware**
  - **using write-through is too expensive so we use writeback**

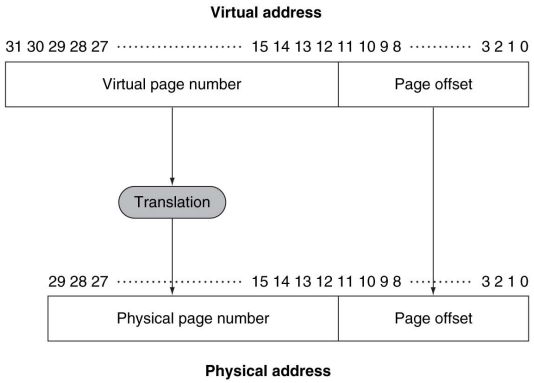


Analogy between cache and VM

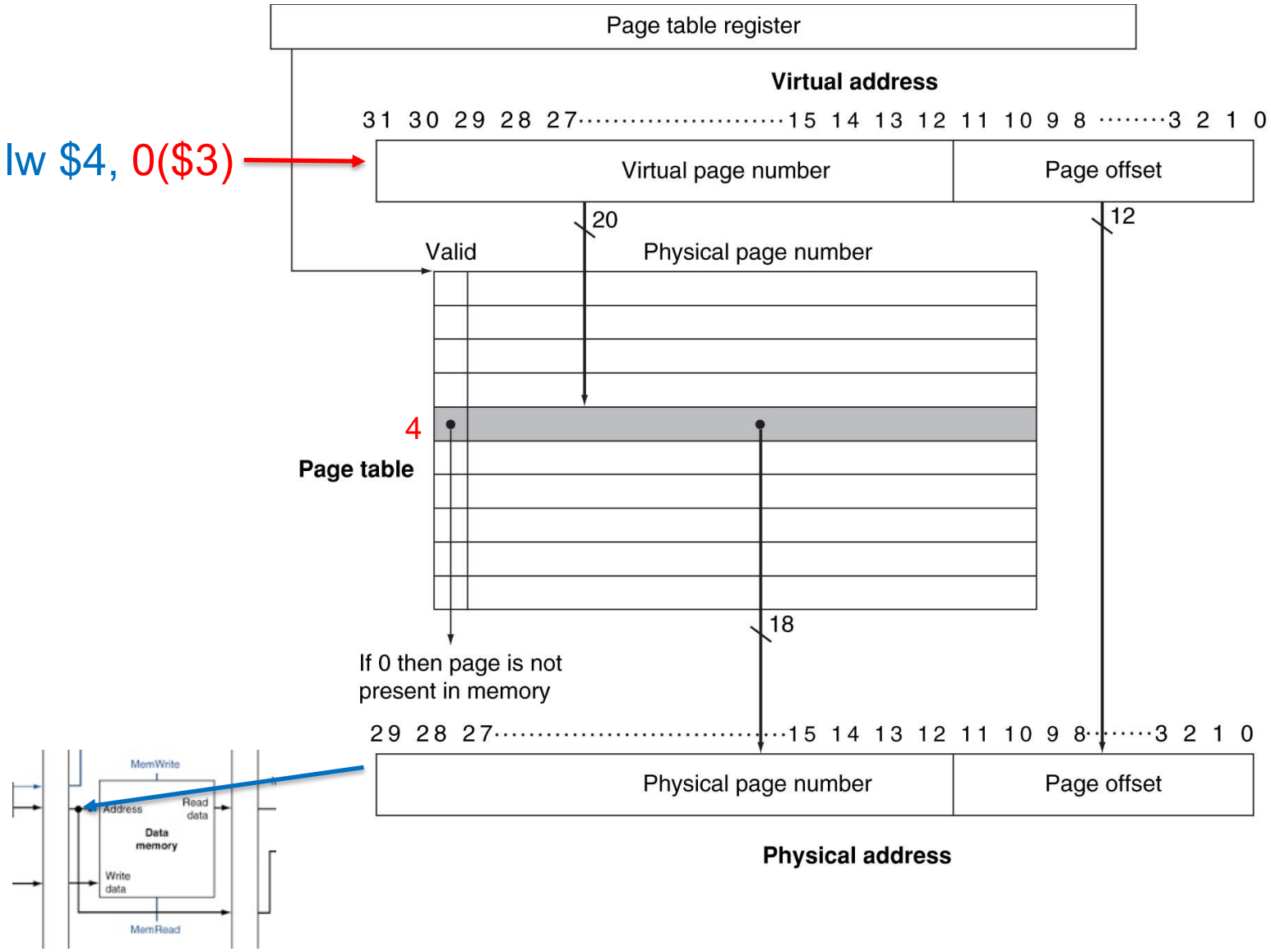
	Cache	VM
Upper level memory	Cache	Main memory
Lower level memory	Main memory	Secondary Storage
Unit of transfer	Block (32~128B)	Page (4~16 KB)
Upper level memory organization	Direct-mapped or set-associative	Fully associative
Replacement scheme	Random	LRU
Write scheme	Write-through or write-back	Write-back
Miss penalty	Relatively low	Relatively large

Page Table : virtual page number → physical page number



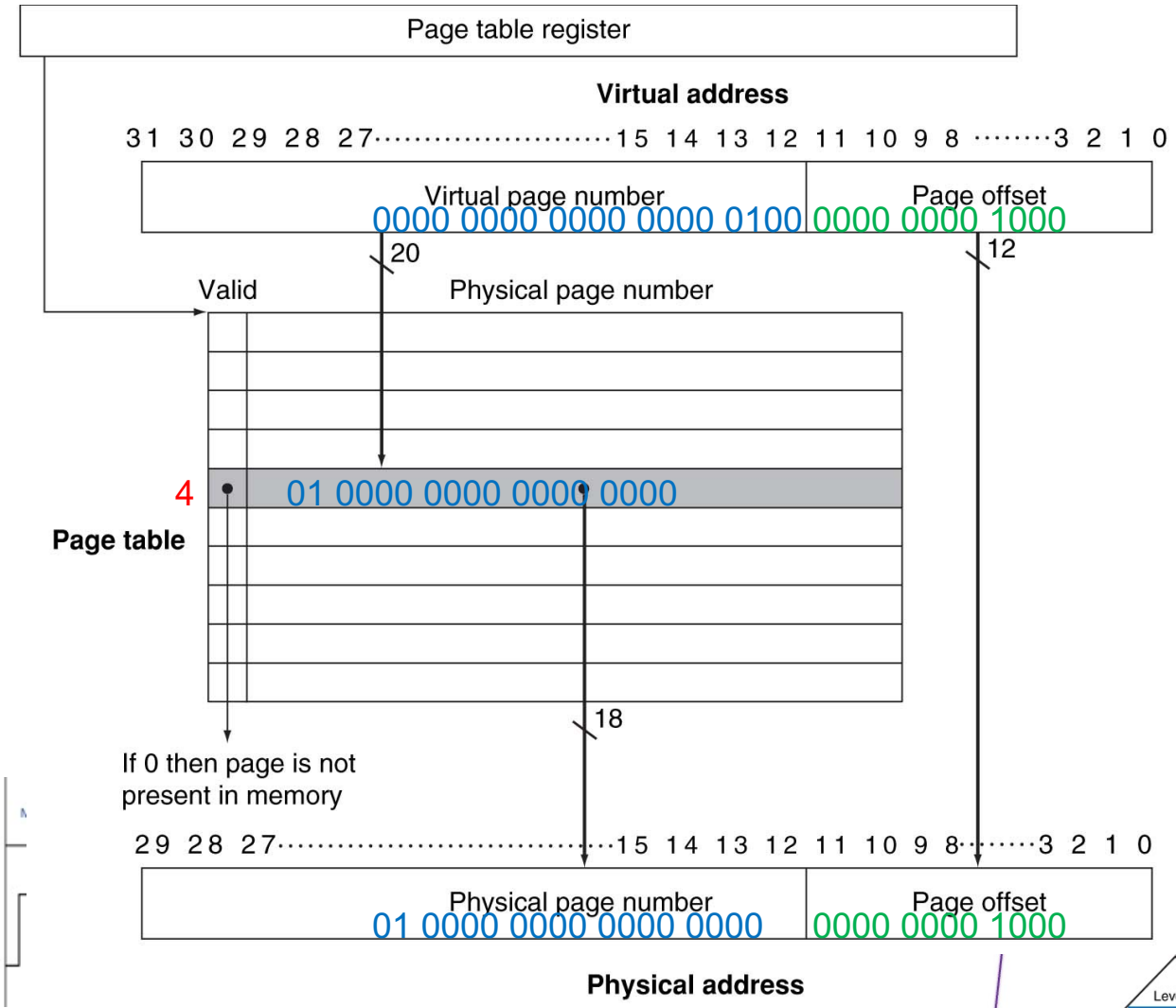


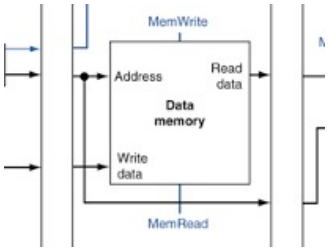
Address translation : virtual address → physical address



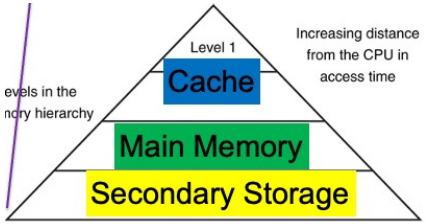
Address translation

lw $4, 0($3)



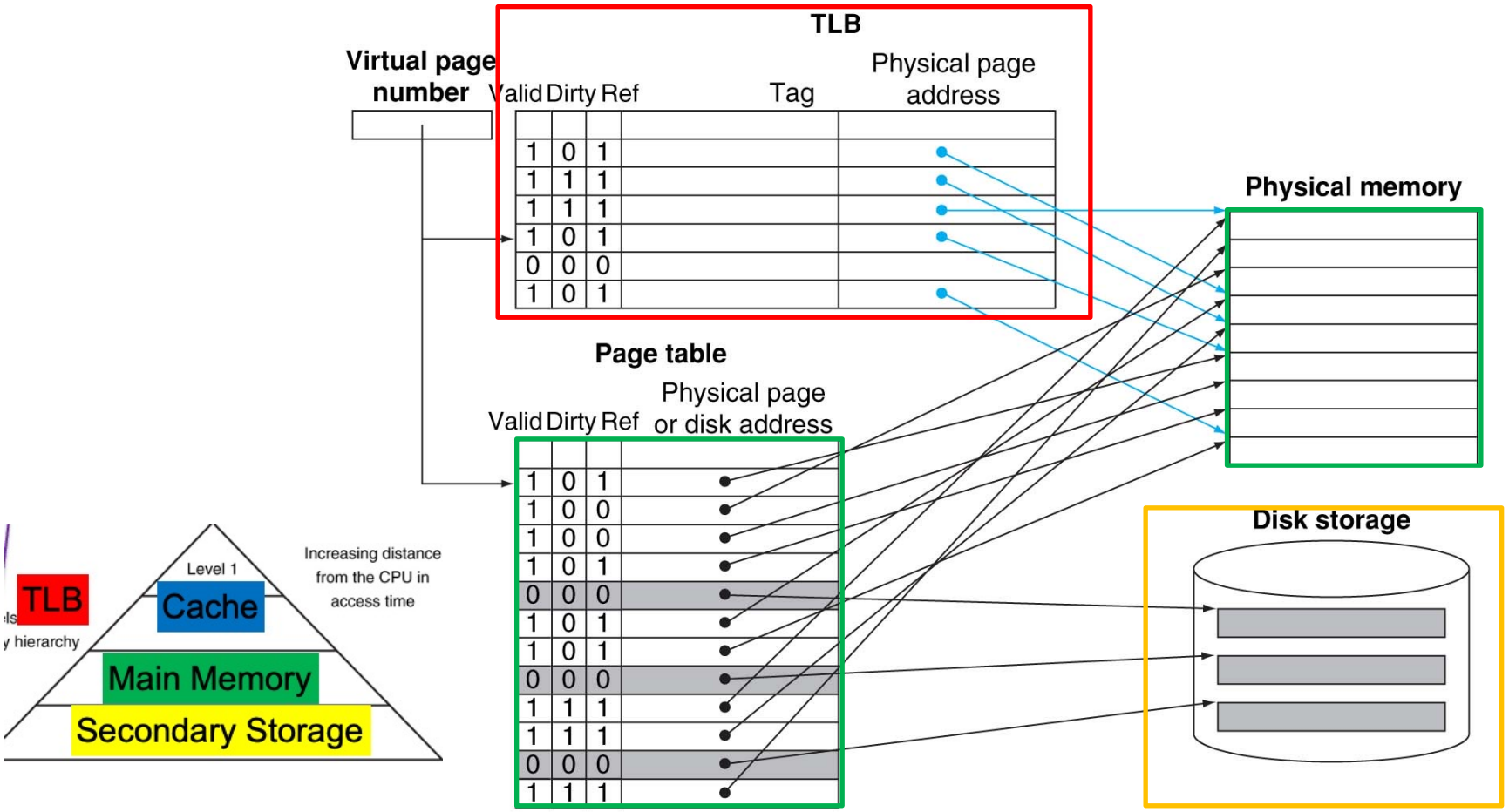


Where is the page table?



Translation Lookaside Buffer (TLB)

- **A cache for address translations: translation lookaside buffer (TLB)**

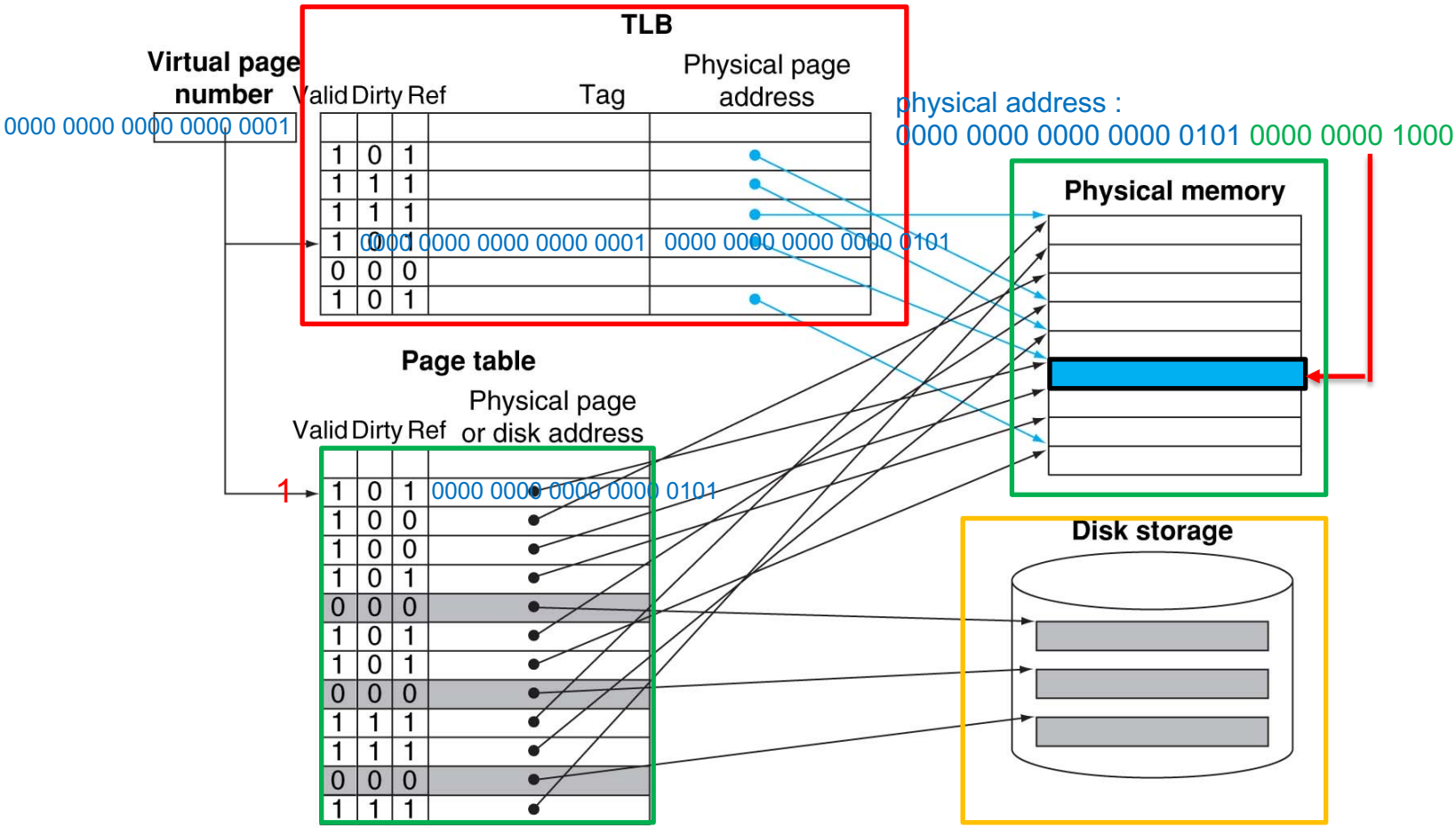


Typical values: 16-512 entries,
miss-rate: 0.01% - 1%
miss-penalty: 10 – 100 cycles

Translation Lookaside Buffer (TLB)

lw $4, 0($3)

virtual address : 0000 0000 0000 0000 0001 0000 0000 1000

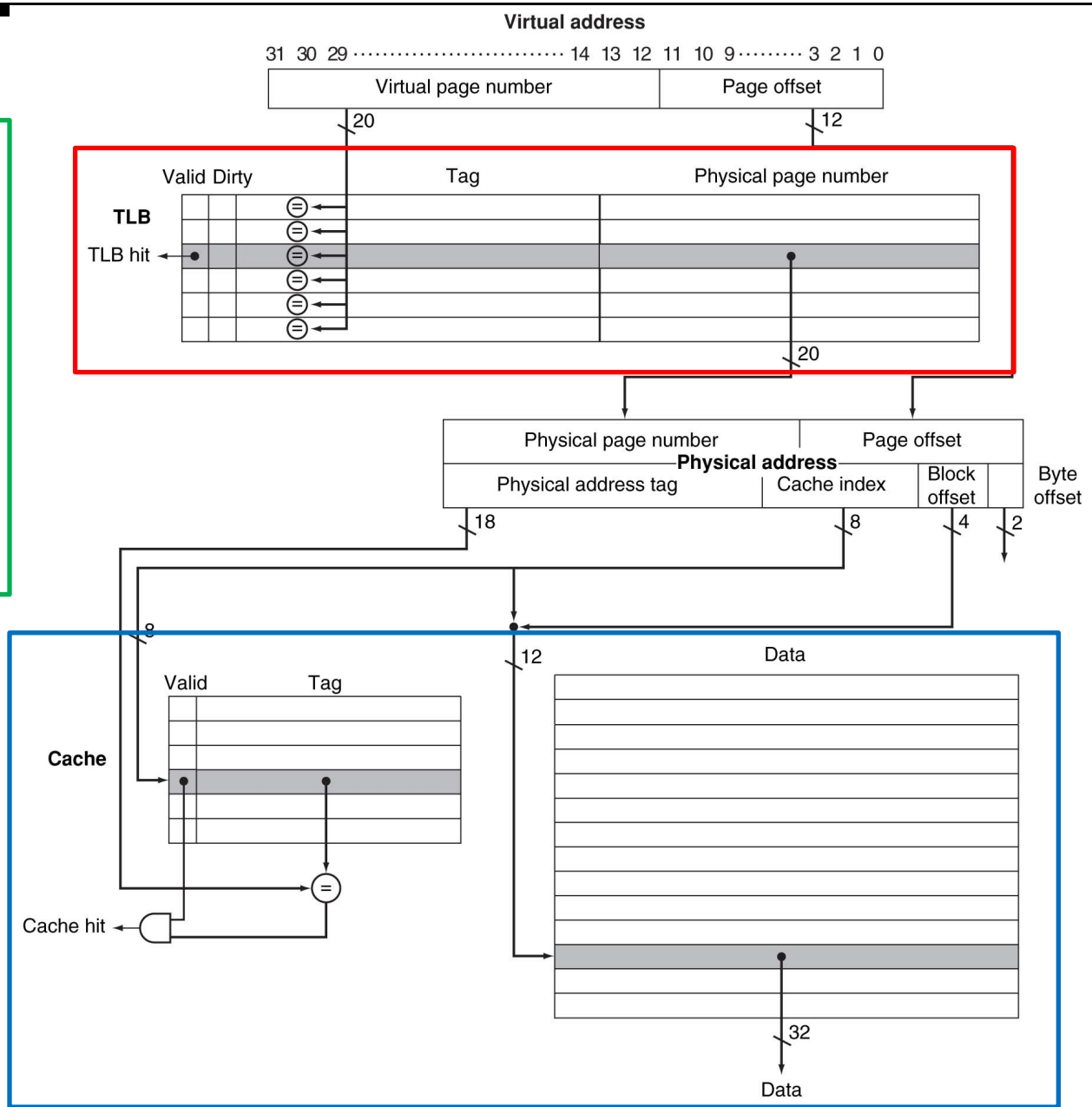


Putting it all together : virtual memory + cache example

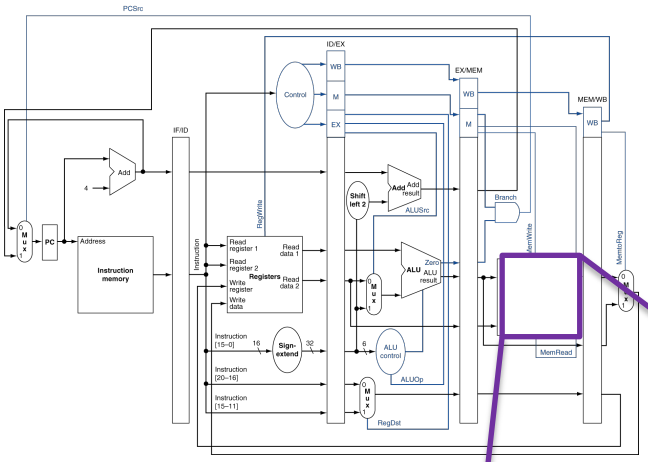
main memory

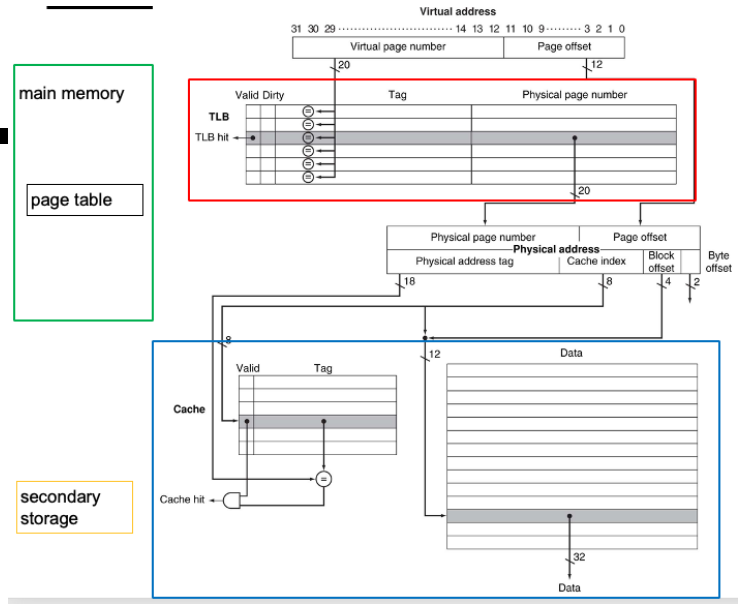
page table

secondary storage

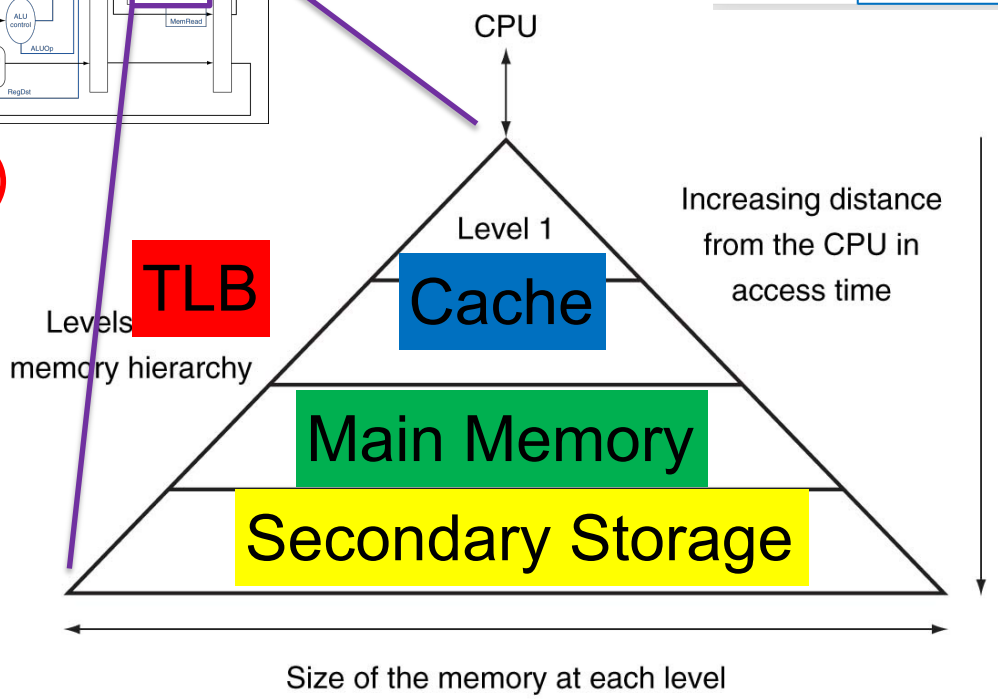


Memory Hierarchy





lw $4, 0($3)



virtual memory + cache example

lw $4, 0($3)

main memory

page table

secondary storage