Operating Systems W10L1 - Memory Management II (ctd)

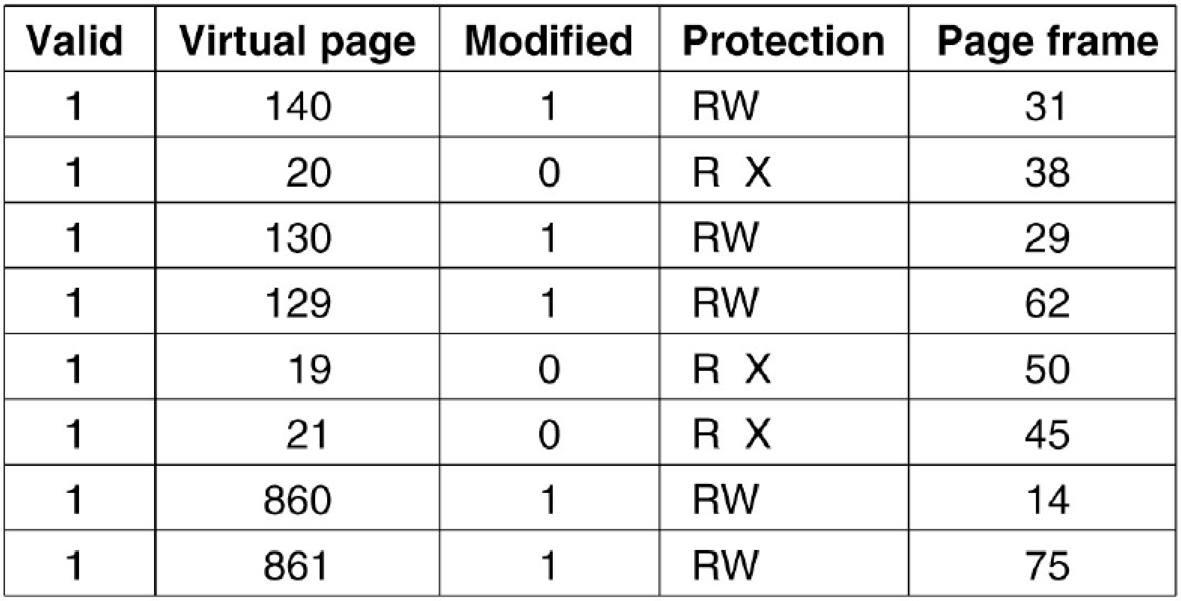
Class	Operating Systems
Created	@Nov 2, 2020 6:50 PM
Materials	
Reviewed	□
Type	Lecture

Speeding Up Virtual Memory

TLB Recap

- Only a fraction of the page tablei s heavily used so we don't need to acccess the entire table. The TLB makes it so we don't have to consistently go to the page table
- It contains page table entries for a small number of pages → A cache for the page table
- See last lecture notes for TLB Hit and TLB Miss diagrams — Additional instructions and explanations in textbook as well
- TLB misses are rare, and that's especially fortunate because if a miss occurs there is an additional memory cost
    - If TLB misses were frequent, we'd be even worse off in terms of performance tha nwithout it
- TLB misses altogether occur more often than page faults

▼ TLB Contents



Valid	Virtual page	Modified	Protection	Page frame
1	140	1	RW	31
1	20	0	R X	38
1	130	1	RW	29
1	129	1	RW	62
1	19	0	R X	50
1	21	0	R X	45
1	860	1	RW	14
1	861	1	RW	75

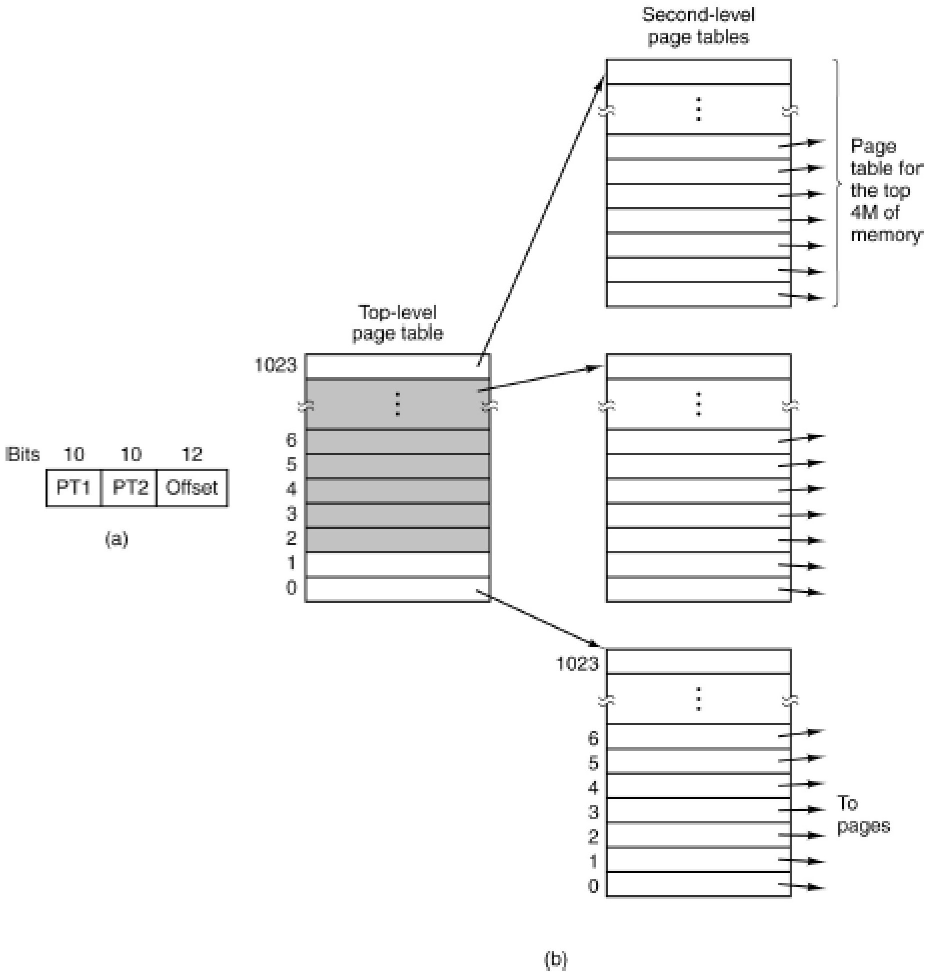
- During a context switch, there are two solutions
  1. Just flush out the TLB entirely
  2. Add a process ID column to the TLB

  The first one could be a security breach since flushing just changes valid bits to 0 and leaves the rest accessible, but the second one is larger and potentially slower

Page Table Size Problem

- We want to be sleective about what is in memory
- We do not know anything about the stack or the heap
  - Stack depends on function calls, heap depends on dynamica allocaiton

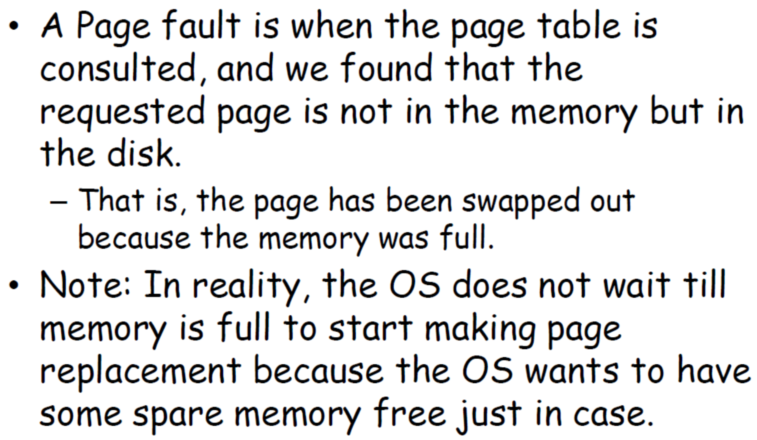
▼ Multi-Level Page Table diagram



1. One solution is to store memory address of smaller page table (start of it) in an "upper level" page table entry
   - This way we only need the small page table in memory
   - Corresponding lower-level page tables are created as needed, thus reducing memory requirement
2. Another solution is an *inverted page table* where there is one entry in the page table per page frame (as opposed to per page)
   - Saves a lot of storage, but makes virtual-to-physical storage much more difficult

Page Faults

▼ What is the definition of a page fault?



- A Page fault is when the page table is consulted, and we found that the requested page is not in the memory but in the disk.
  - That is, the page has been swapped out because the memory was full.
- Note: In reality, the OS does not wait till memory is full to start making page replacement because the OS wants to have some spare memory free just in case.