Schedulers

Turnaround time = completion time – arrival time

Response time = first run time – arrival time

FIFO – bad turnaround, suffers from Convoy effect

SJF – alright if all commands arrive at first time. Convoy effect arises if not

STCF (Shortest to complete first) – Good turnaround, not the best response

RR – great response, not the best turnaround

MLFQ and Lottery others to consider

MLFQ: Rule 1: If Priority(A) > Priority(B), A runs (B doesn’t). If Priority(A) = Priority(B), A & B run in RR. When a job enters the system, it is placed at the highest priority (the topmost queue). If job uses entire time slice while running, its priority is reduced. If job gives up CPU before time slice is up, stays at same priority level

Preemptive – Scheduler can decide to run another job when one arrives, continuing the current one later

Traps and Interrupts

-Returns from the kernel require a different instruction, which both undoes what the trap did (re-stores saved state), and changes privileged mode back to user mode

-Retry only happens in some cases (e.g., TLB miss); in other cases (e.g., syscall trap), the return must return to the instruction after the trapping instruction

Os Paging

-OS has to set up page table base register on startup and switch PTBR on context switches. Must also change page tables on memory allocation and free. Also monitors accesses to be able to perform page replacement.

Virtual Address = [VPN, OFFSET] where VPN is offset into page table and offset is the page offset. Use VPN to offset into page table, grab the PTE. PTE consists of valid bits and the PFN – [page base address, valid bits] Real address = [PFN, OFFSET]

Read, write, execute, etc. PFN, valid bits, and prot bits are in both TLB and PTE

Paging allows avoiding external fragmentation and flexible address usage. It is generally not faster than segmentation and is more complex. It can also lead to internal fragmentation.

Software managed TLB – software does the translations, vs a hardware TLB hardware does translations

TLB Valid bit says translation is accurate, Page table valid bit says the page is in use.

Threads share page table base register and address space

Don’t share instruction pointer or common registers, or stack. They can see each others stacks.