Date No

CPU Scheduling - Short term Scheduling (SOS) lecture 05

- Cpu Scheduling used in multiprogramming systems to utilize the cpu usage in positive manner.

- have 2 types of burst in Scheduling.
  - Cpu burst
  - I/O burst

- cpu burst - The time took the process to being execute in the processor.

- I/O burst - When the cpu is whaiting for I/O for further execution.

cpu burst → (I/O) I/O burst → (ready que) cpu burst

- Start with cpu burst and end with also cpu burst.

- have 2 types of bound mesurese.
  - I/O bound
  - CPU bound.

- CPU ~~I/O~~ bound - time took to Complete a task determind with Speed of cpu. have few but long cpu burst
- I/O bound - time took to complete I/O operation. have Short cpu burst

Cpu Skheduler.

- this will select an process from ready que and alocate to cpu.
- Cpu Schedular will select a new process only in,
  1. process switches from running state to whoiting state (I/o whait).
  2. Switches from running to ready (Interrupt)
  3. Whaiting to ready (after I/o completion).
  4. Terminates (end of a process).

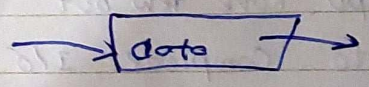
1 and 4 → must select a new process.
2 and 3 → can select any.

- Cpu Skheduling has 2 types
  - preemptive scheduling.
  - non-preemptive Scheduling.

- Preemptive - has a timer, each prosses given a time to execute in cpu. (newdays)

- non-preemptive - Cpu only ocleased by the time of termination or to move to Whaiting que. no timer,

- Preemptive Scheduling can cause race conditions. many prosses access the same data at given time



Date No

Dispatcher.

- this component give the control of the cpu to the Scheduler that select the current process. (short term scheduler or cpu scheduler.

  - The time took to stop a process start another called dispatch latency.

  - Dispatcher involves,
    - When switching context
    - running user application
    - When restarting a program.

Using below points we can mesure the algorithem.

- cpu utilization - max
- Throughput - max
- Turnaround time - min
- Whe'ting time - min
- Repond time - min

refer 5.10 for more.

There are 3 types of cpu scheduling algorithems

- first come, first served (FCFS)
- Shortes job first (SJF)
- Round Robin (RR)

02) first come first served (FCFS)

- Which process coming first will be served
- if the burst time was short average

Whaiting time will be reduced.

Convoy effect - Short processes behind long proces

to this solution SJF algorithem came.

03) Shortest Job first, (SJF)

- SJF gives minimum whaiting time
- In here first have to find the cpu burst.
- In preemptive caled Shortest-remaining-time-first.

Cpu burst can be predected using exopential averagh formula.

$$\tau_{n+1} = \alpha t_n + (1 - \alpha)\tau_n$$

- if there is no recent histor $\alpha = 1$;
- if there it will be the Value.
- if we dont care $\alpha = 0$.

Date No

(03) Round Robin algorithem (RR)

- each process gets a small unit of cpu time. (time quantem) (10 - 100 milliseconds).

- in round robin performance depend on a time quantem.

- if q is large → FCFS
- if q is small → have to do context switch

- We want quantem to be larg.
- good for interactive systems.

(04) Priority Scheduling.

- based on the priority cpu will allocated and each process given a ~~time~~ number

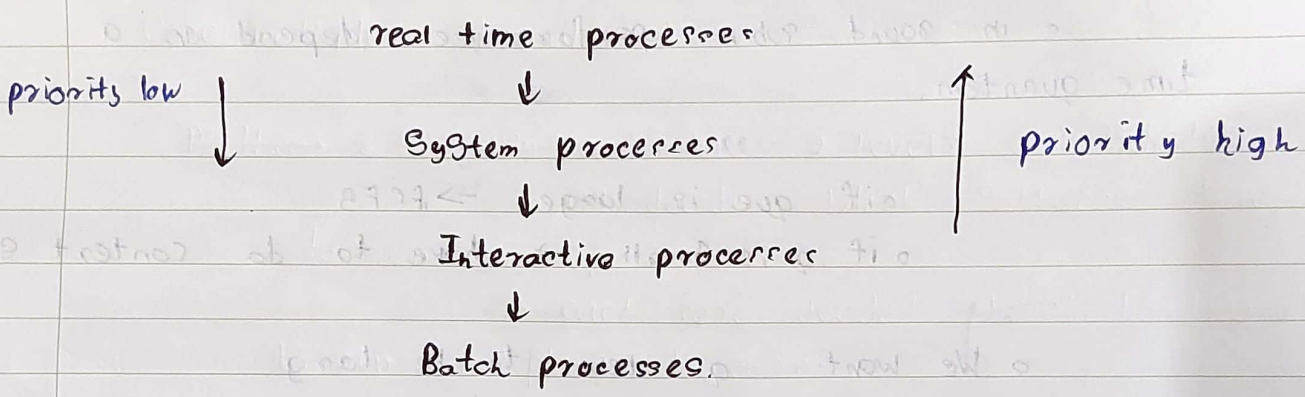
- Smallest priority → highest integer and vise versa.

- SJF is priority Scheduling method.

Starvation - low priority processes may not execut
Solutions (aging) - increase the priority of the process.

(05) Multilevel que.

- We can have seperate ques for each priority. prioritization can do based on process type.



(06) Multivelev feedback ques.

- here processes can move between ques.

refere 5.30 to more.

- by this method low priority ques can be moved to higher priority que.

Thread Scheduling.

there a 2 types of threads.

- user-level threads
- kernel-level threads.

- if the threads support then threads will be scheduled not the processes.

Date No

- Thread library will schedule user-level threads to light weigh processes. (LWP). also known as process contention scope (PCS)

- System Contention scope(SCS) will decide which the kernel thread schedul into cpu. this

~~Pthread Scheduling~~,

Multiprocessor Scheduling.

- this is used when a system has more than a single processor.
- Multiprocessors can be,
  - Multicore cpu's
  - Multi threaded core's
  - NUMA system
  - Heterogeneous multiprocesing

Symmetric multiprocessing (SMP)

processer schedule it self. has a common ready que. also has a private que of threads

refer S. 98 for more., S. 99

Multithreaded Multicore System.

- each core has a thread. if one thread has not enough space it will switch to another thread.
- Chip multi threading (CMT) - assings each Core multiple threads.

~~Multicor~~
- this System has 2 levels of Scheduling.
  - OS decides which software thread to run on cpu.
  - each core decided which hardware to run on physical core.

Load balancing,
- how the work load is distributed among processors. 2 ways to do it.
  - Push migration
  - Pull migration.

5.42

Processor afinity.
- When thread running on one processor the cache of that processer information of processor stored by that threads.
- Load balancing effect afinity. 2 types of affinity.
  - soft afinity
  - hard afinity

5.43

Date No

Real time Cpu Scheduling.

- have 2 types of real time systems. 5.45
  - Soft real time
  - hard real time.

Event latency.

- the time took to event occurse to being serviced. 2 types of latencys. 5.46
  - Interrupt latency
  - Dispatch latency.

Algorithm Evaluation.

- Deterministic modeling.
  - We have predfined work load and test it on each algorithm. (performance).
- Queueing model.
  - Compute the throughput, Utilization, waiting estimate the cpu burst and arrival time distribution.
- little's formula.
  - Valid for any algorithm and for any arrival distribution.
- Simulation.
  - to get more accurate result we can use this.
- implimentation