CMP 321 OS 22/01/24

Process control block

|  |
| --- |
| Process state |
| Process |
| Process content |
| Process register |
| CPU Schedule |
| List of open files |

Content Switch

When CPU switches from one process to another process, the system saves the state of the old process and load the saved state for the new process.

State of process

A process is a sequence of actions, performed by executing a sequence of instructions.

Process creation

* When a process is added, the OS operation builds a data structure, and allocates address to the memory. So we can create a process and terminate it as well

Process termination

A process terminates when it finishes executing its final statement, and asks the OS to delete it.

Cooperative process

The concurrent processes executing in the operating system might either be independent or cooperating. A process is independent if it cannot be affected by other processes in execution. In other words, any process that does not share data with another process is independent. The process is cooperating if it can affect or be affected by other processes during execution. Clearly, any process that shares data with other process (A dependent process) is called a cooperating process.

Reasons for cooperating processes

* Information sharing
* Computational speed up
* Convenience
* Modularity

Challenges of cooperating process

* Race condition
* Critical region
* Mutual exclusion

Assignment

* Explain the following
* Race condition
* Critical region
* Mutual exclusion

Process scheduling

The aim of process scheduling is to assign process to be executed by the processor overtime in a way that meets system objectives. There are 4 types of scheduling;

1. Long term
2. Medium term
3. Short term
4. Input-output scheduling

(NON)Pre-emptive policies

In a preemptive scheme, short term scheduler may remove a process from running state. E.g. in a non-pre-emptive scheme, once a process is to the processor it cannot be removed or terminated until it is executed except it encounters an **I/O wait.**

Algorithm for Process Schedule

* First come first serve (FCFS/FIFO): in this policy, execute the oldest process in the ready queue. It simply assigns the processor to process the first task in the ready queue
* Shortest job first(next) (SJF/SJN): It is an approach to reduce the bias. It uses a non-preemptive policy in which the process with shortest time is executed first.
* Round Robin: In order to reduce the penalty of SJF, it assigns equal time to all the processes and executes them as so. In RR, each process has a specified quantum of time to run. But if the process runs more than the fixed time, it is interrupted and sent back to the ready state.
* Shortest remaining time (SRT): In SRT, it uses preemptive policy. However, it compares the remaining time of the present task to the remaining task provided the time is shorter than the present task. In summary, the shorter ongoing task is executed. E.g.
* Race condition: is an undesirable situation that occurs when a device or system tries to perform 2 or more operations at a time.
* Mutual exclusion: is a frequently used method for synchronizing processes or threads that want to access some shared resources.
* Critical region: refers to a section of code or data structure that must be accessed exclusively by one method or thread at a time.

Concurrency is described as the ability of a system to handle multiple tasks or processes simultaneously, it allows multiple processes to make progress and overlap in execution even if the underlined hardware is only able to execute one task at a time through techniques like multi-tasking.

* Process priority
* Resource availability
* Concurrency and parallelism
* System load
* I/O bound vs CPU bound
* Processor architecture
* Algorithm efficiency
* External factors
* Measurement and benchmarking issues.

PROBLEMS OF CONCURENCY

* Priority management

Context switching: is the process where the CPU saves the state of one process and loses the state of another process to enable multi-tasking.

Advantages

* Multi-tasking

Disadvantages

* Memory overload
* Loss of process data

Process scheduling: refers to the process where an operating system manages the allocation of system resources to processes or threads.

Long scheduling: process move from the queue to the memory for execution

Short-scheduling: selects which process goes next.

Medium scheduling: swaps the processes between main memory and storage.

I/O scheduling: puts process in wait until it receives a directive from the I/O

Queues in an operating system:

Difference between long and short term – long from new while short from ready

Scheduling criteria:

FIFO, SJF, SRT, round Robbin: is a preemptive policy which assigns a dedicated amount of time to each task. Provided that the tasks have not surpassed their time limit, the processes will continue to execute in a circular motion.

Process control block is a data structure which manages all the information about a process. Process creation is the action of the OS to start a new process.