

CSE-316

**FOUR LEVEL PRIORITY DISPATCHER**

K18PD-G2-65

Student Name: Lov Kumar

Student ID : 11804637

Email Address: lovkumar33@gmail.com

GitHub Link: <https://github.com/lovkumar33/OS-Assignment>

Submitted To :

Amandeep Kaur

**Q35**. Write a four level priority process dispatcher operating within the constraints of finite available memory and I/O resources. Use different dispatchers and memory allocation.



**Solution : -**

The process scheduling is the activity of the process manager that handles the removal of the running process from the CPU and the selection of another process on the basis of a particular strategy.

Process scheduling is an essential part of a Multiprogramming operating systems. Such operating systems allow more than one process to be loaded into the executable memory at a time and the loaded process shares the CPU using time multiplexing.

**Dispatcher :**

Another component that is involved in the CPU-scheduling function is the dispatcher, which is the module that gives control of the CPU to the process selected by the short-term scheduler. It receives control in kernel mode as the result of an interrupt or system call. The functions of a dispatcher mop the following:

* [Context switches](https://en.wikipedia.org/wiki/Context_switch), in which the dispatcher saves the [state](https://en.wikipedia.org/wiki/State_(computer_science)) (also known as [context](https://en.wikipedia.org/wiki/Context_(computing))) of the [process](https://en.wikipedia.org/wiki/Process_(computing)) or [thread](https://en.wikipedia.org/wiki/Thread_(computing)) that was previously running; the dispatcher then loads the initial or previously saved state of the new process.
* Switching to user mode.
* Jumping to the proper location in the user program to restart that program indicated by its new state.

The dispatcher should be as fast as possible, since it is invoked during every process switch. During the context switches, the processor is virtually idle for a fraction of time, thus unnecessary context switches should be avoided. The time it takes for the dispatcher to stop one process and start another is known as the dispatch latency.

**Problem Explained :**

The problem of scheduling of resources like operators and machines in manufacturing

systems has been studied over the last four decades. Research work has been carried out on the theory of scheduling and its practical applications. Methods based on branch and bound technique, priority dispatching rules, simulated annealing, tabu search, genetic algorithm, artificial neural networks, expert systems, linear programming, integer programming, graph theory, petri nets and lagrangian relaxation techniques etc., have been proposed. Objectives such as the minimization of lead time and work in process inventory, and maximisation of throughput, resource utilisation and timely delivery have been considered.

The scheduling studies into the following two

categories :

* Theoretical research dealing with optimisation procedures.
* Experimental research dealing with dispatching rules.

The theoretical research has aimed at developing mathematical models and optimal or

sub optimal algorithms. The theoretical results have not been widely used in the industries due to the high computational complexity of the scheduling problem. The experimental research has been primarily concerned with priority dispatching rules and heuristics that can efficiently solve the scheduling problem.

**Processes :**

Processes on HOST are simulated by the dispatcher creating a new process for each dispatched process. This process is a generic process (supplied as process — source: sigtrap.c) that can be used for any priority process. It actually runs itself at very low priority, sleeping for one-second periods and displaying the following:

1. A message displaying the process ID when the process starts;
2. A regular message every second the process is executed; and
3. A message when the process is Suspended, Continued, or Terminated.

The process will terminate of its own accord after 20 seconds if it is not terminated by your dispatcher. The process prints out using a randomly generated color scheme for each unique process, so that individual "slices" of processes can be easily distinguishable. Use this process rather than your own.

The life cycle of a process is as follows:

The process is submitted to the dispatcher input queues via an initial process list that designates the arrival time, priority, processor time required (in seconds), memory block size, and other resources requested.

A process is "ready-to-run" when it has "arrived" and all required resources are available.

Any pending Real-Time jobs are submitted for execution on a first-come-first-served basis.

If enough resources and memory are available for a lower priority User process, the process is transferred to the appropriate priority queue within the feedback dispatcher unit, and the remaining resource indicators (memory list and i/o devices) updated.

When a job is started (fork and exec("process",...)), the dispatcher will display the job parameters (Process ID, priority, processor time remaining (in seconds), memory location and block size, and resources requested) before performing the exec.

A Real-Time process is allowed to run until its time has expired when the dispatcher kills it by sending a SIGINT signal to it.

A low priority User job is allowed to run for one dispatcher tick (one second) before it is suspended (SIGTSTP) or terminated (SIGINT) if its time has expired. If suspended, its priority level is lowered (if possible) and it is requeued on the appropriate priority queue as shown in Figures P2.1 and P2.3. To retain synchronization of output between your dispatcher and the child process, your dispatcher should wait for the process to respond to a SIGTSTP or SIGINT signal before continuing ( waitpid(p->pid**,** &status**,** WUNTRACED**)**). To match the performance sequence indicated in the comparison of scheduling policies (see Figure 9.5), the User job should not be suspended and moved to a lower priority level unless another process is waiting to be (re)started.

Provided no higher-priority Real-Time jobs are pending in the submission queue, the highest priority pending process in the feedback queues is started or restarted (SIGCONT).

When a process is terminated, the resources it used are returned to the dispatcher for reallocation to further processes.

When there are no more processes in the dispatch list, the input queues and the feedback queues, the dispatcher exits.

**Dispatch List :**

The Dispatch List is the list of processes to be processed by the dispatcher. The list is contained in a text file that is specified on the command line. That is,

>hostd dispatchlist

Each line of the list describes one process with the following data as a *"comma-space*" delimited list:

<arrival time>, <priority>, <processor time>, <Mbytes>, <#printers>, <#scanners>, <#modems>, <#CDs>

Thus,

12, 0, 1, 64, 0, 0, 0, 0

12, 1, 2, 128, 1, 0, 0, 1

13, 3, 6, 128, 1, 0, 1, 2

would indicate the following:

**1st Job:** Arrival at time 12, priority 0 (Real-Time), requiring 1 second of processor time and 64 Mbytes memory — no I/O resources required.

**2nd Job:** Arrival at time 12, priority 1 (high priority User job), requiring2 seconds of processor time, 128 Mbytes of memory, 1 printer, and 1 CD drive.

**3rd Job:** Arrival at time 13, priority 3 (lowest priority User job), requiring 6 seconds of processor time, 128 Mbytes of memory, 1 printer, 1 modem, and 2 CD drives.



