UNIVERSITY of INFORMATION TECHNOLOGY

Operating Systems Fundamentals

Faculty of Computer Science

UNIVERSITY of INFORMATION TECHNOLOGY

Uniprocessor Scheduling

9

1. Type of Processor Scheduling
2. Scheduling Algorithms I
3. Scheduling Algorithms II

Faculty of Computer Science

Learning Objectives

- To explain the differences among long-term, medium-term, and short-term scheduling
- To describe the performance of different scheduling policies.
- To understand the scheduling technique used in traditional UNIX.

Processor Scheduling

UNIVERSITY of INFORMATION TECHNOLOGY

- Aim is to assign processes to be executed by the processor in a way that meets system objectives, such as
  - response time,
  - throughput, and
  - processor efficiency
- Broken down into three separate functions:

Table 9.1
Types of Scheduling

Long-term scheduling	The decision to add to the pool of processes to be executed
Medium-term scheduling	The decision to add to the number of processes that are partially or fully in main memory
Short-term scheduling	The decision as to which available process will be executed by the processor
I/O scheduling	The decision as to which process's pending I/O request shall be handled by an available I/O device

Scheduling and Process State Transitions

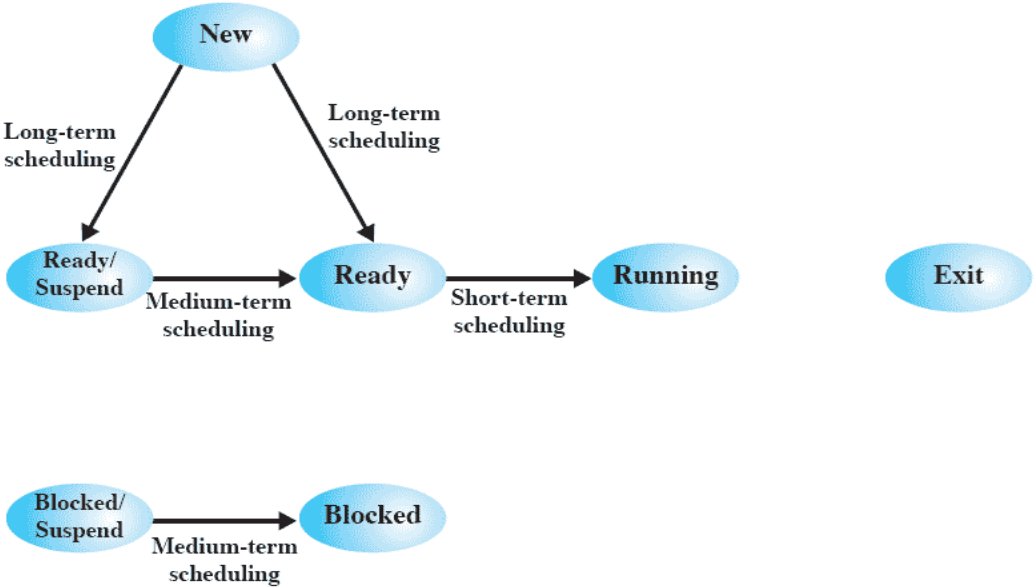
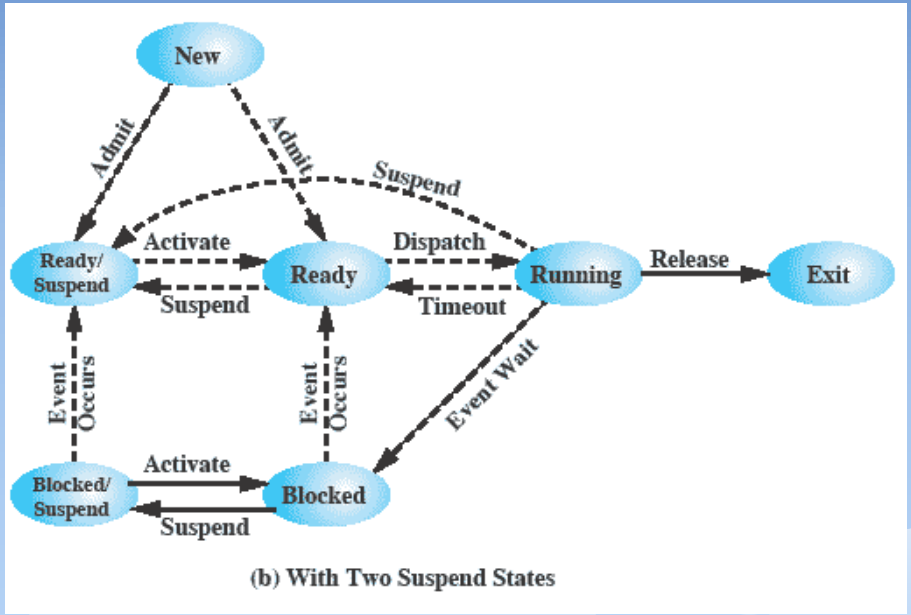


Figure 9.1 Scheduling and Process State Transitions

(Referencing figure 3.9b)



(b) With Two Suspend States

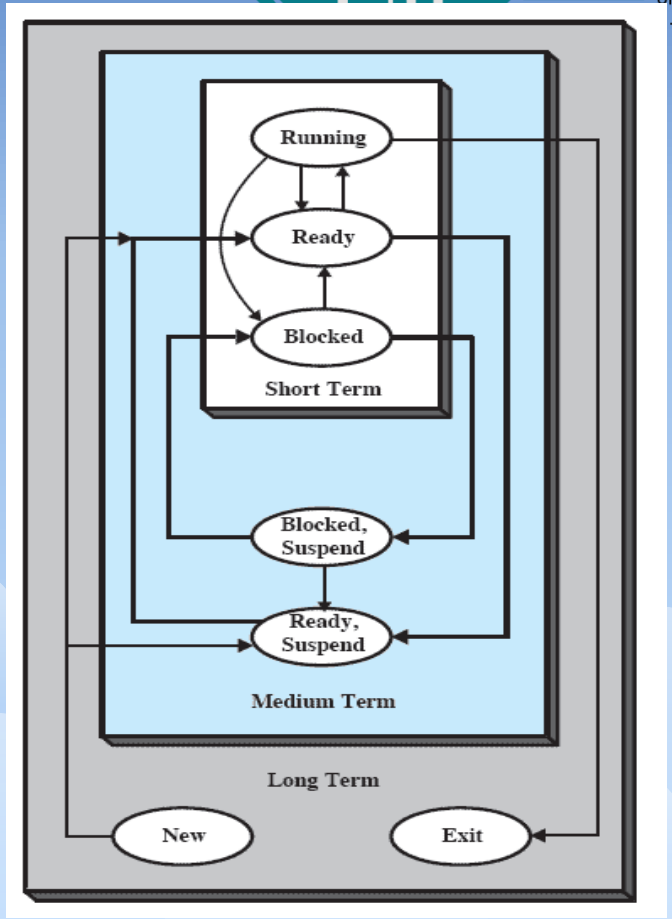


Figure 9.2 Nesting of

Scheduling Functions

Queuing Diagram

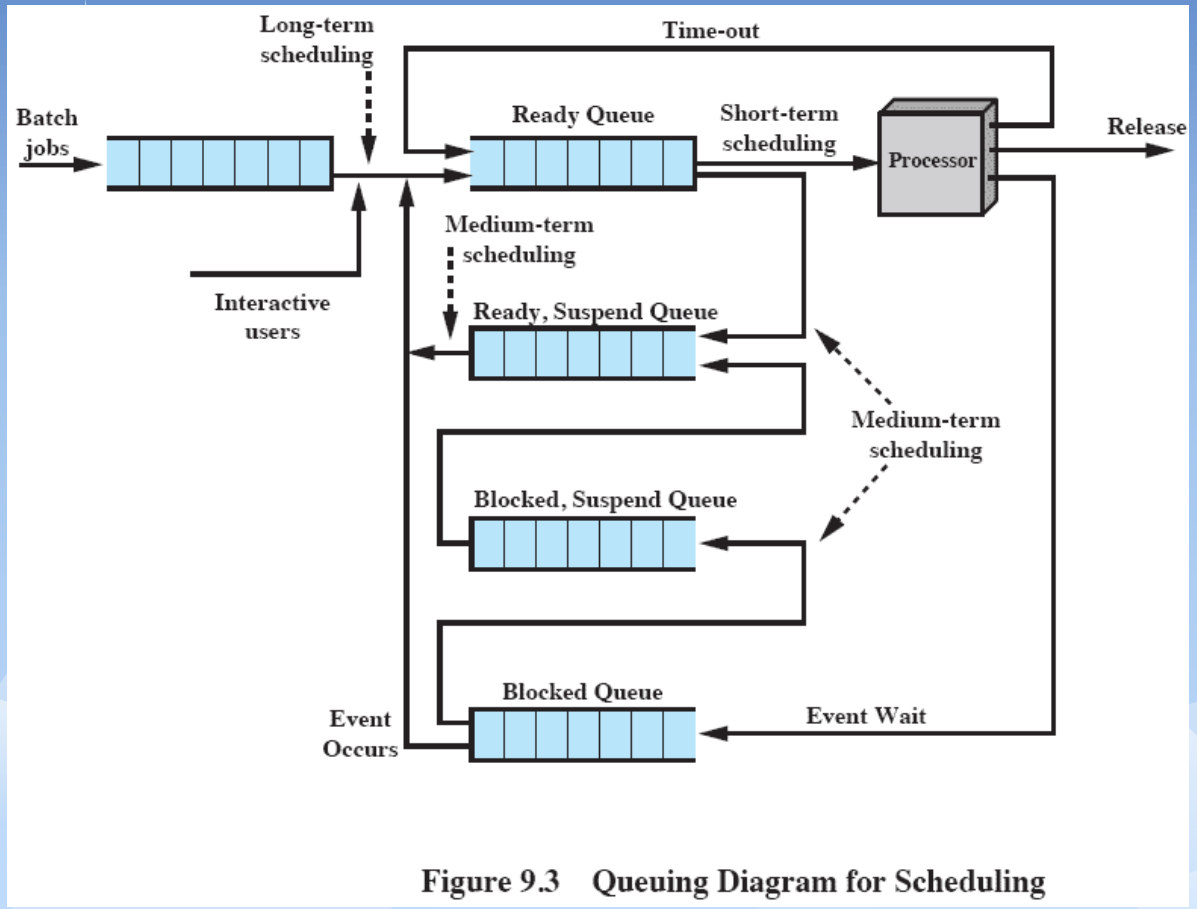


Figure 9.3 Queuing Diagram for Scheduling

Long-Term Scheduler

UNIVERSITY of INFORMATION TECHNOLOGY

- Determines which programs are admitted to the system for processing
  - May be first-come-first-served
  - Or according to criteria such as priority, I/O requirements or expected execution time

- Controls the degree of multiprogramming
  - the more processes that are created, the smaller the percentage of time that each process can be executed
  - may limit to provide satisfactory service to the current set of processes

Medium-Term Scheduling

UNIVERSITY of INFORMATION TECHNOLOGY

- Medium-term scheduling is part of the swapping function.
- **Swapping-in** decisions are based on the need to manage the degree of multiprogramming, considers the memory requirements of the swapped-out processes.

Short-Term Scheduling

UNIVERSITY of INFORMATION TECHNOLOGY

- Known as the **dispatcher.**
- Executes most frequently.
- Makes the fine-grained decision of which process to execute next
- Invoked when an event occurs that may lead to the blocking of the current process or that may provide an opportunity to preempt a currently running process in favor of another.

Examples:

- Clock interrupts
- I/O interrupts
- Operating system calls
- Signals (e.g., semaphores)

Short Term Scheduling Criteria

UNIVERSITY of INFORMATION TECHNOLOGY

- Main objective is to allocate processor time to optimize certain aspects of system behaviour.
- A set of criteria is needed to evaluate the scheduling policy

User-oriented criteria

- relate to the behavior of the system as perceived by the individual user or process (such as ***response time*** in an interactive system)
- important on virtually all systems

System-oriented criteria

- focus in on effective and efficient utilization of the processor ***(throughput : rate at which processes are completed***)
- generally of minor importance on single-user systems

Short-Term Scheduling Criteria: Performance

UNIVERSITY of INFORMATION TECHNOLOGY

examples:
- response time
- throughput

example:
- predictability

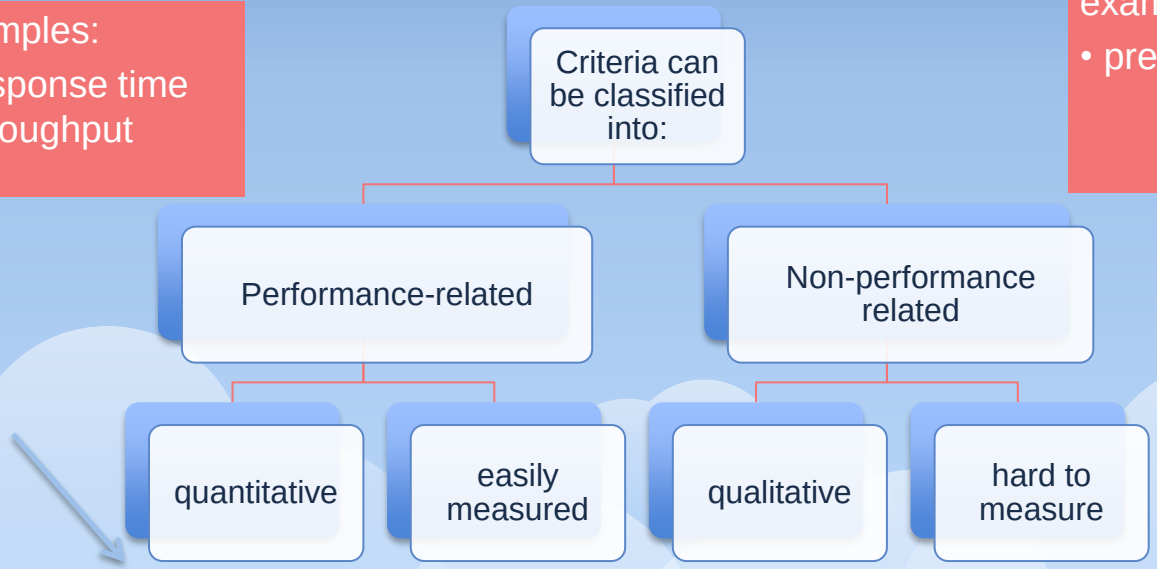


Table 9.2 Scheduling Criteria

User Oriented, Performance Related

Turnaround time This is the interval of time between the submission of a process and its completion. Includes actual execution time plus time spent waiting for resources, including the processor. This is an appropriate measure for a batch job.

Response time For an interactive process, this is the time from the submission of a request until the response begins to be received. Often a process can begin producing some output to the user while continuing to process the request. Thus, this is a better measure than turnaround time from the user's point of view. The scheduling discipline should attempt to achieve low response time and to maximize the number of interactive users receiving acceptable response time.

Deadlines When process completion deadlines can be specified, the scheduling discipline should subordinate other goals to that of maximizing the percentage of deadlines met.

User Oriented, Other

Predictability A given job should run in about the same amount of time and at about the same cost regardless of the load on the system. A wide variation in response time or turnaround time is distracting to users. It may signal a wide swing in system workloads or the need for system tuning to cure instabilities.

System Oriented, Performance Related

Throughput The scheduling policy should attempt to maximize the number of processes completed per unit of time. This is a measure of how much work is being performed. This clearly depends on the average length of a process but is also influenced by the scheduling policy, which may affect utilization.

Processor utilization This is the percentage of time that the processor is busy. For an expensive shared system, this is a significant criterion. In single-user systems and in some other systems, such as real-time systems, this criterion is less important than some of the others.

System Oriented, Other

Fairness In the absence of guidance from the user or other system-supplied guidance, processes should be treated the same, and no process should suffer starvation.

Enforcing priorities When processes are assigned priorities, the scheduling policy should favor higher-priority processes.

Balancing resources The scheduling policy should keep the resources of the system busy. Processes that will underutilize stressed resources should be favored. This criterion also involves medium-term and long-term scheduling.

UNIVERSITY of INFORMATION TECHNOLOGY

Table 9.2 Scheduling Criteria

References

- Operating Systems: “**Internals and Design Principles**” Seventh Edition By William Stallings, 2012
- http://www.cs.uah.edu/~weisskop/Notes490/Chapter05alt-OS7e.ppt
- http://www.cs.uah.edu/~weisskop/Notes490/Chapter06alt-OS7e.ppt