Scheduling

Motivation and definition of terms

1. Ordering resource assignments (CPU time, network access, ...)
2. Time scales of scheduling
   - Short-term
   - Medium-term
   - Long-term
   - Historical

Purpose of scheduling

- Stallings, William
  - Principles of Concurrent and Dis-
  - guage and Standard Libraries;

- ISO/IEC 8652:201x (E)
  - to permit admittance of new requests or for load-balancing, ...

Defnition of terms

Performance scheduling

- Requested resource times
  - Predictability
  - Minimize the duration of any single request and the variance across requests.

- Throughput
  - Maximize the number of requests completed.

- CPU busy time
  - Minimize the number of times the CPU is idle.

- Waiting time
  - Optimize for short waiting times for all requests.

- Turnaround time
  - Optimize for short turnaround times for all requests.

- Requested resource times
  - Minimize the time from given a new request to a new request's end.

- Minimize the number of times a process is interrupted.

- Optimize for short waiting times for all requests.