## Scheduling Criteria

- Throughput # of threads that complete per unit time
  # jobs/time (Higher is better)
- Turnaround time time for each thread to complete
  Tfinish Tstart (Lower is better)
- Response time time from request to first response ()
  i.e. time between waiting to ready transition and ready to running transition
  Tresponse Trequest (Lower is better)
- → Above criteria are affected by secondary criteria
  - CPU utilization %CPU fraction of time CPU doing productive work
  - Waiting time Avg(Twait) time each thread waits in the ready queue

## How to balance criteria?

- Batch systems (supercomputers)
  strive for job throughput and turnaround time
- Interactive systems (personal computers) strive to minimize response time for interactive jobs
  - However, in practice, users prefer predictable response time over faster but highly variable response time
  - Often optimized for an average response time