

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

Two kinds of scheduling algorithm

- **Non-preemptive scheduling** (good for batch systems)
once the CPU has been allocated to a thread, it keeps the CPU until it terminates
- **Preemptive scheduling** (good for interactive systems)
CPU can be taken from a running thread and allocated to another