Planning Considerations for Job Scheduling in HPC Clusters Notes

- Two types of components
  1. servers
  11. networking interconnects
- Two classes of clusters
  1. high-throughput computing clusters } maximize throughput
     - large # of nodes, low-end interconnects
  11. high performance computing clusters } minimize communication overhead
     - faster and higher performance interconnects than high throughput clusters

- Resource Management System
  1. Resource Manager ⇄ queue info & node load & resource availability
  11. job scheduler

- wait time - time spent in queue ← priority, load, availability
- turnaround time = wait time + execution time
- response time - how fast user gets response from system after submission
- resource utilization - useful work performed
- sys. throughput - # of jobs completed / unit time
- job width - # of processors requested
- mapping - how subtasks are assigned to processors
- the fair-share strategy utilizes historical data to adjust job priority

- Challenging to max utilization with good mean response time

- Scheduling algorithms
  - I. Time-Sharing - divide processor time into slots
  - II. Space-Sharing - give resource single job until execution is complete
    - FCFS, FIFO, SJF, LJF, RR
      - max utilization at the expense of turnaround time

- Advance Reservation
  - use user predicted execution time to reserve resources

- Backfill - improvement on space sharing
  - takes an advance reservation schedule and places small jobs into the gaps in schedule
  - needs runtime estimate