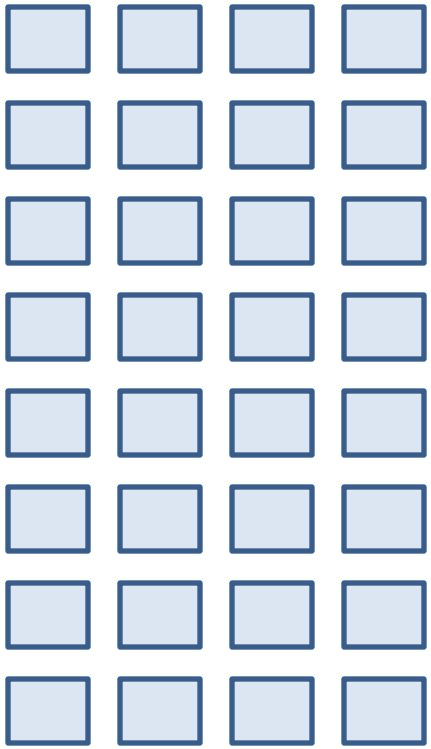
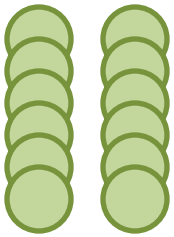
Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud

Gunho Lee (UC Berkeley)
Byung-Gon Chun (Yahoo! Research)
Randy H. Katz (UC Berkeley)

We have resources and jobs

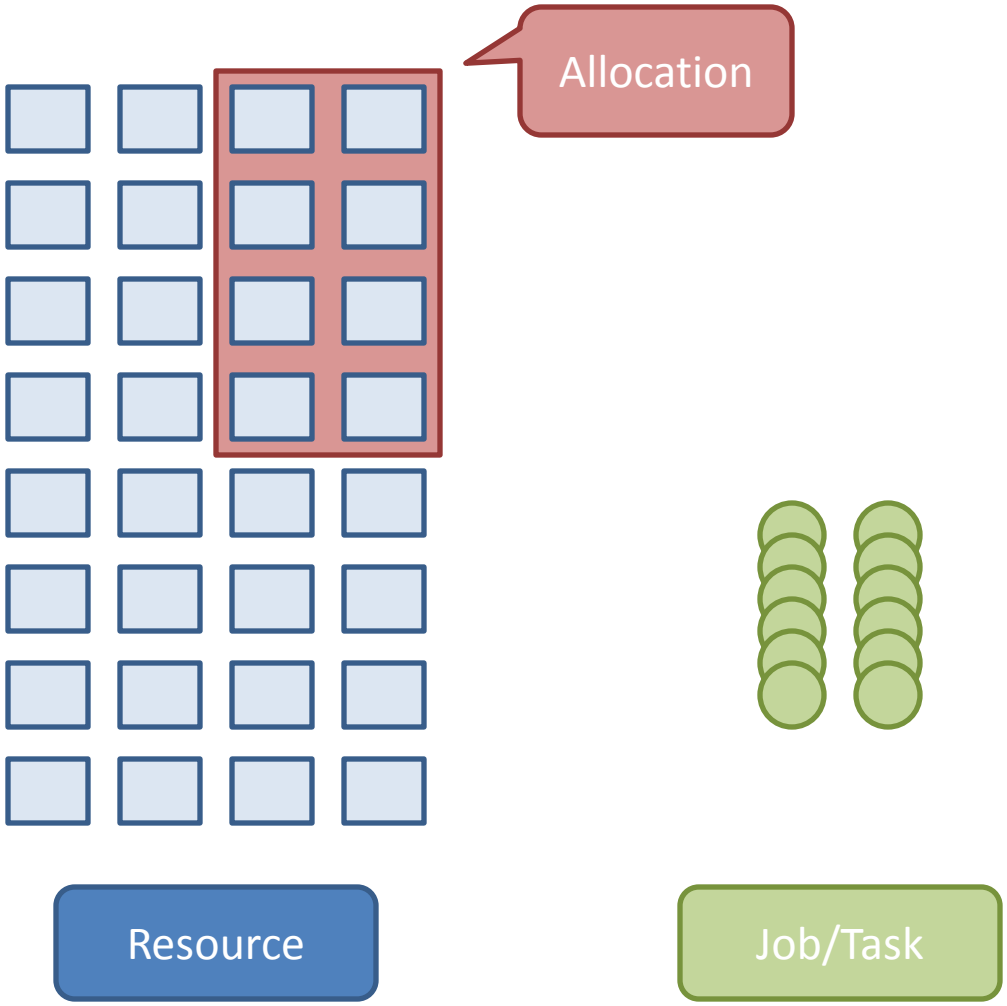


Resource

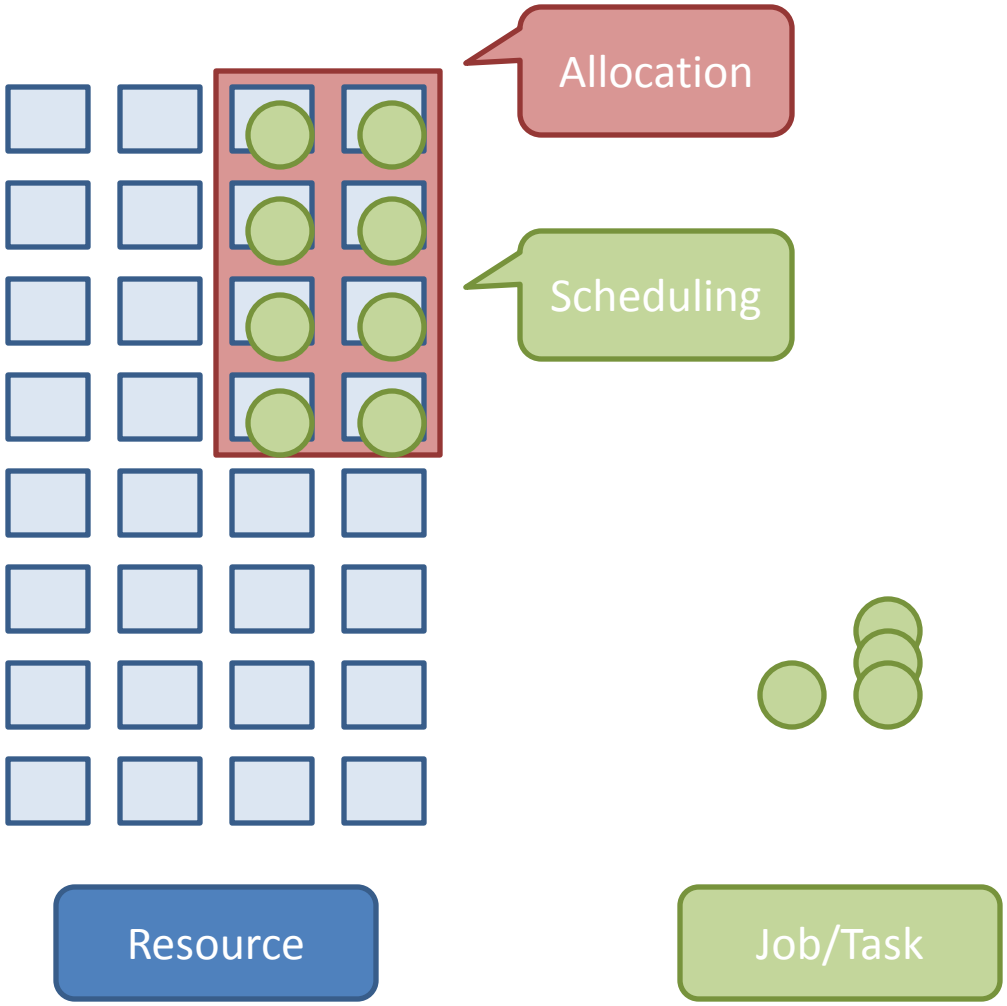


Job/Task

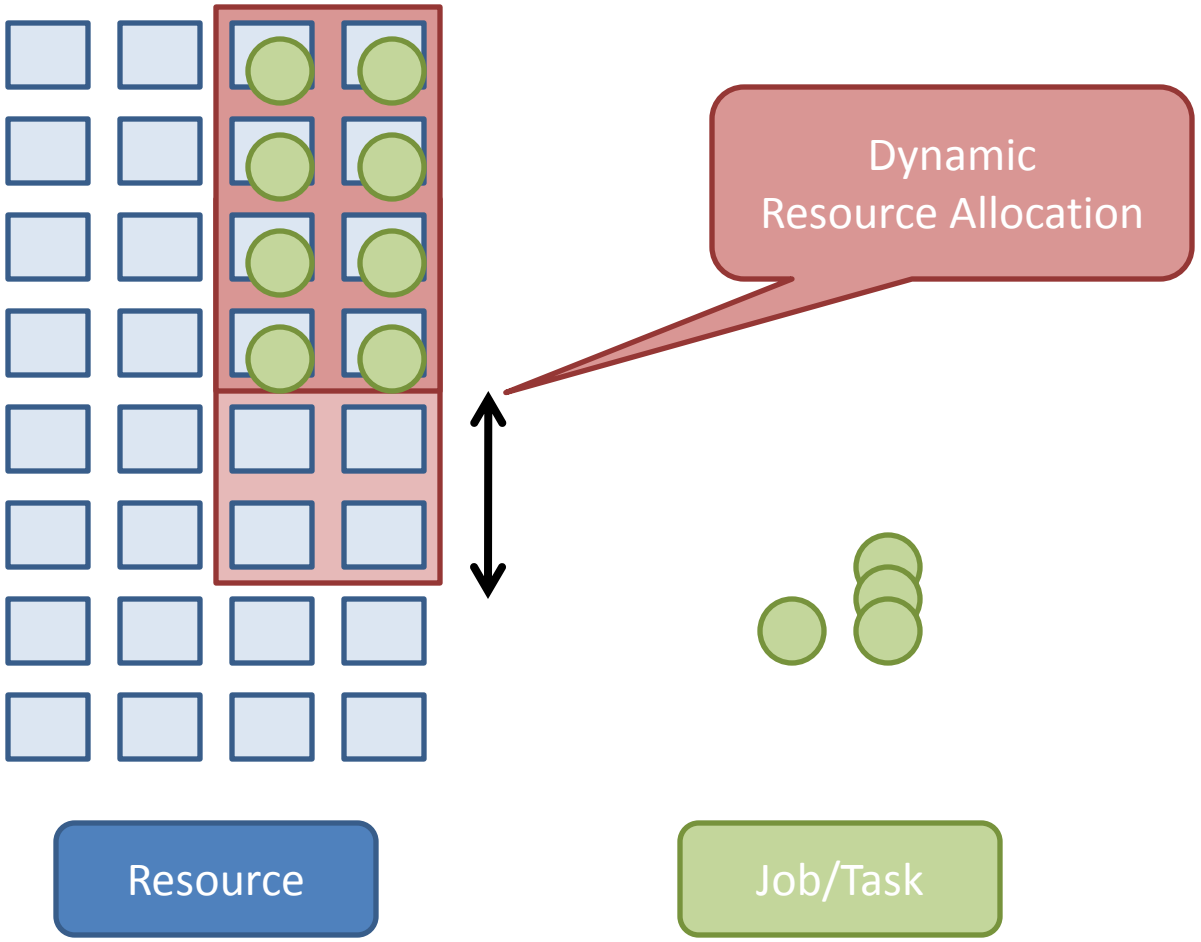
Allocate resources (slots)



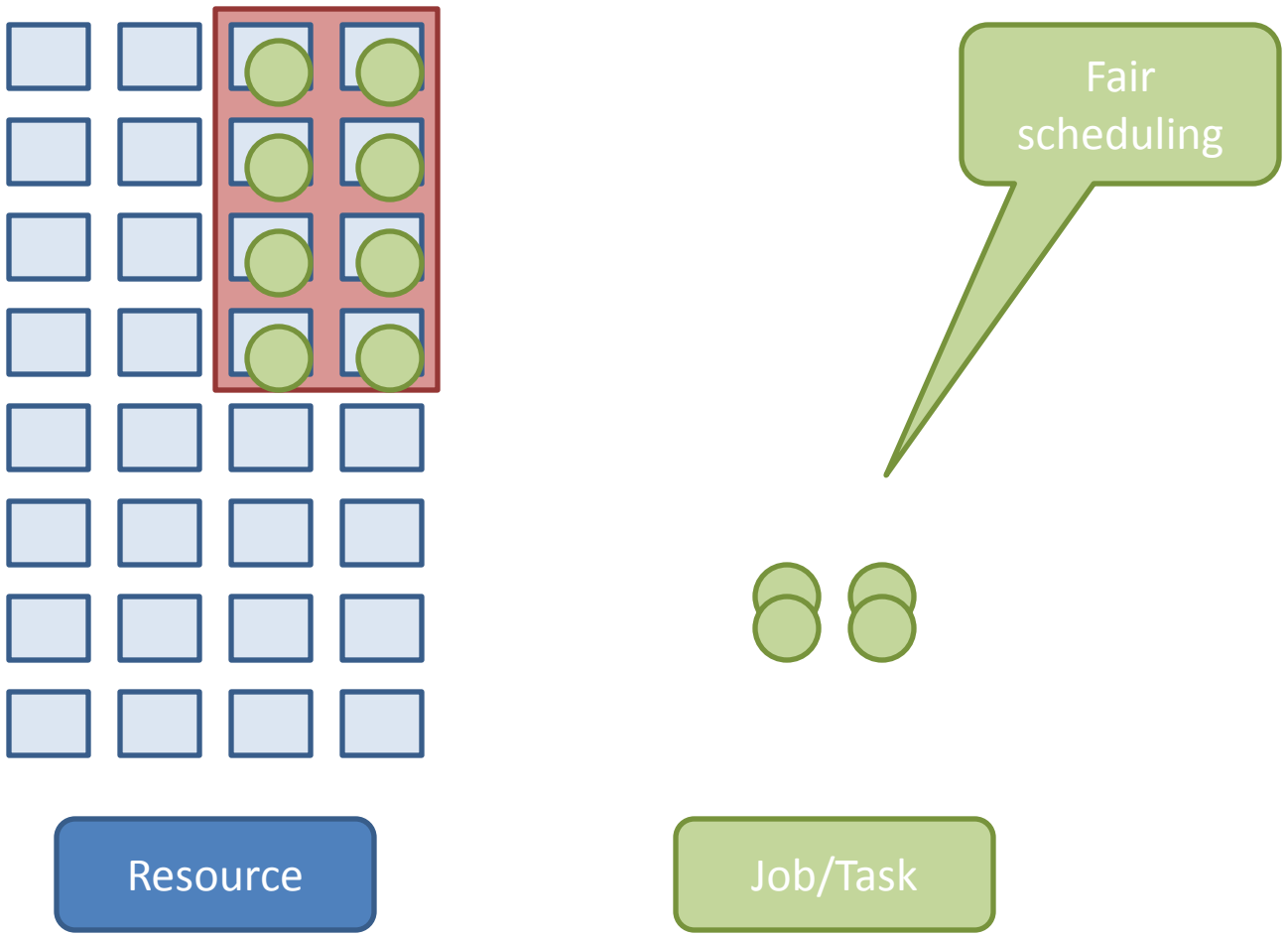
Then schedule jobs/tasks on them



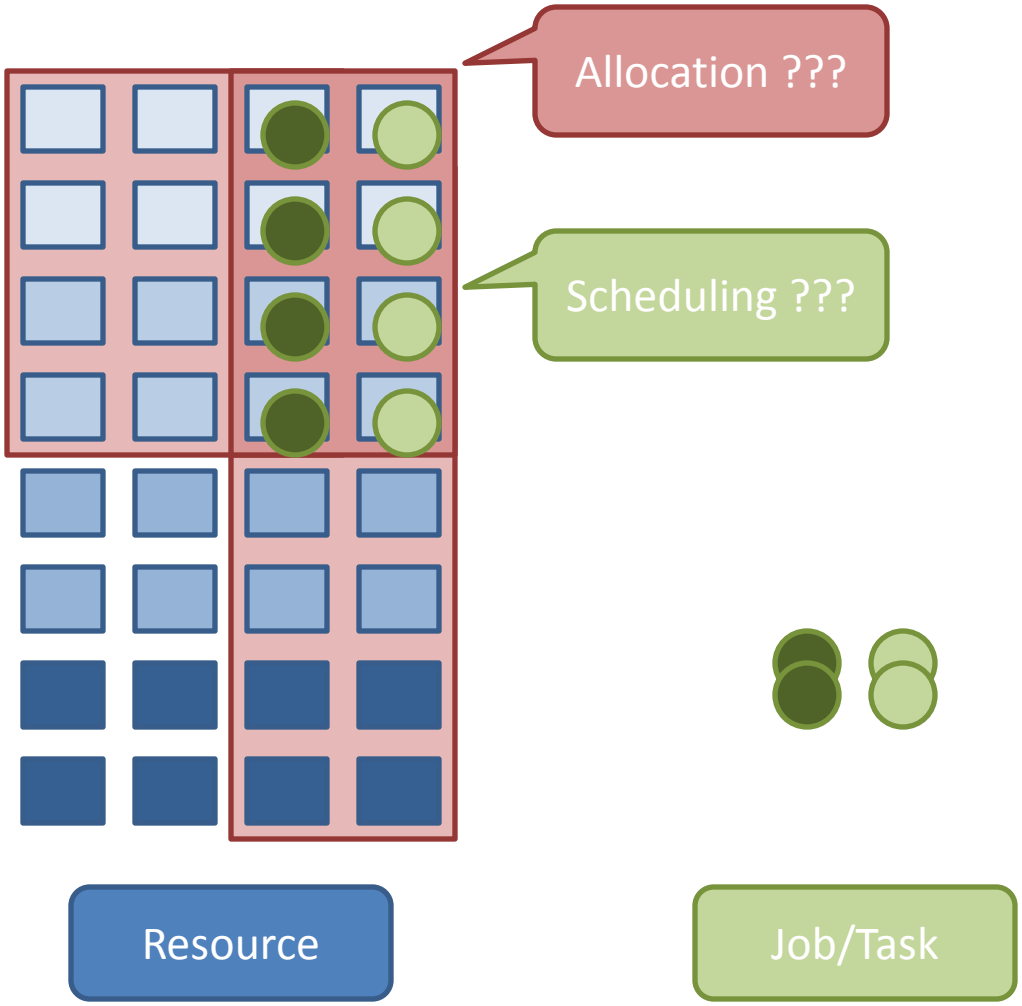
Goal 1. Minimize the cluster size while providing good performance



Goal 2. Provide each job with "fair share" of resources



Heterogeneity makes the problem more complex



Our Approach

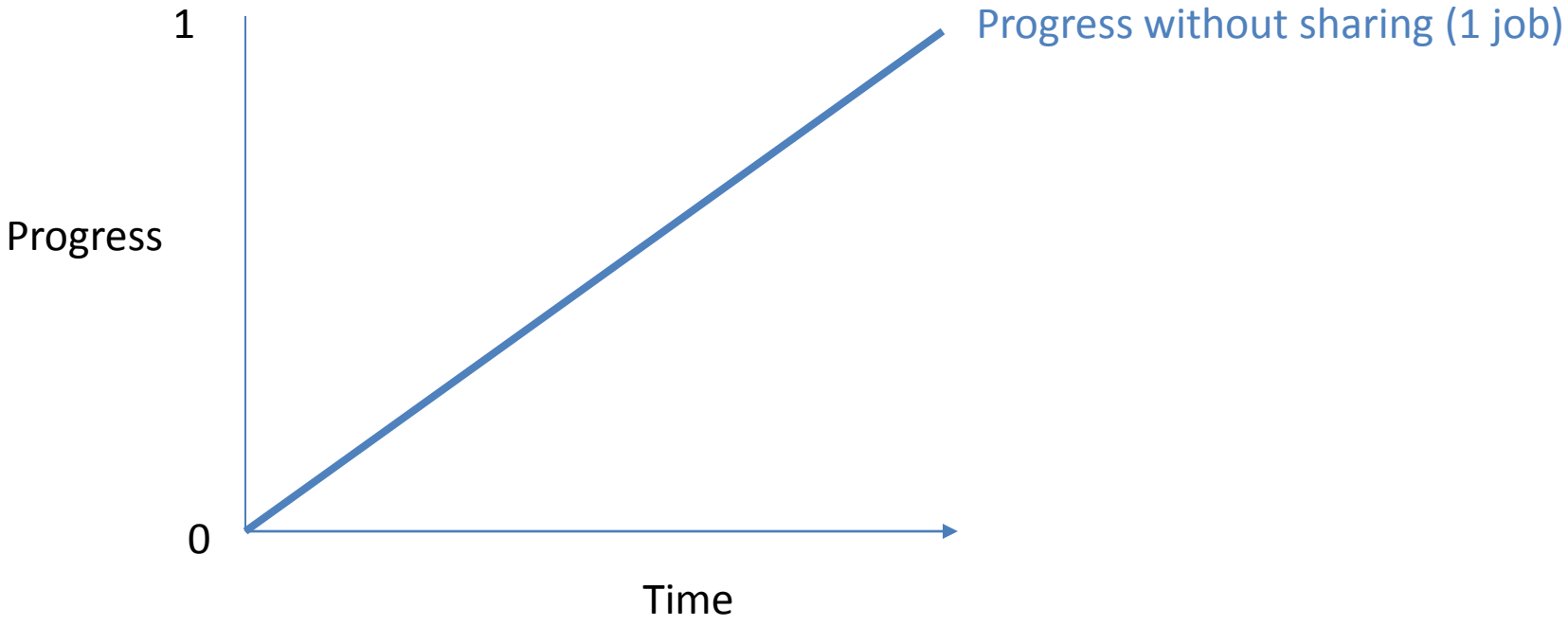
- Consider *Job Affinity* to match more suitable resources to jobs
- Redefine a share metric to provide fairness

- Allocation
  - Core Nodes + Accelerator Nodes
- Scheduling
  - Progress Share

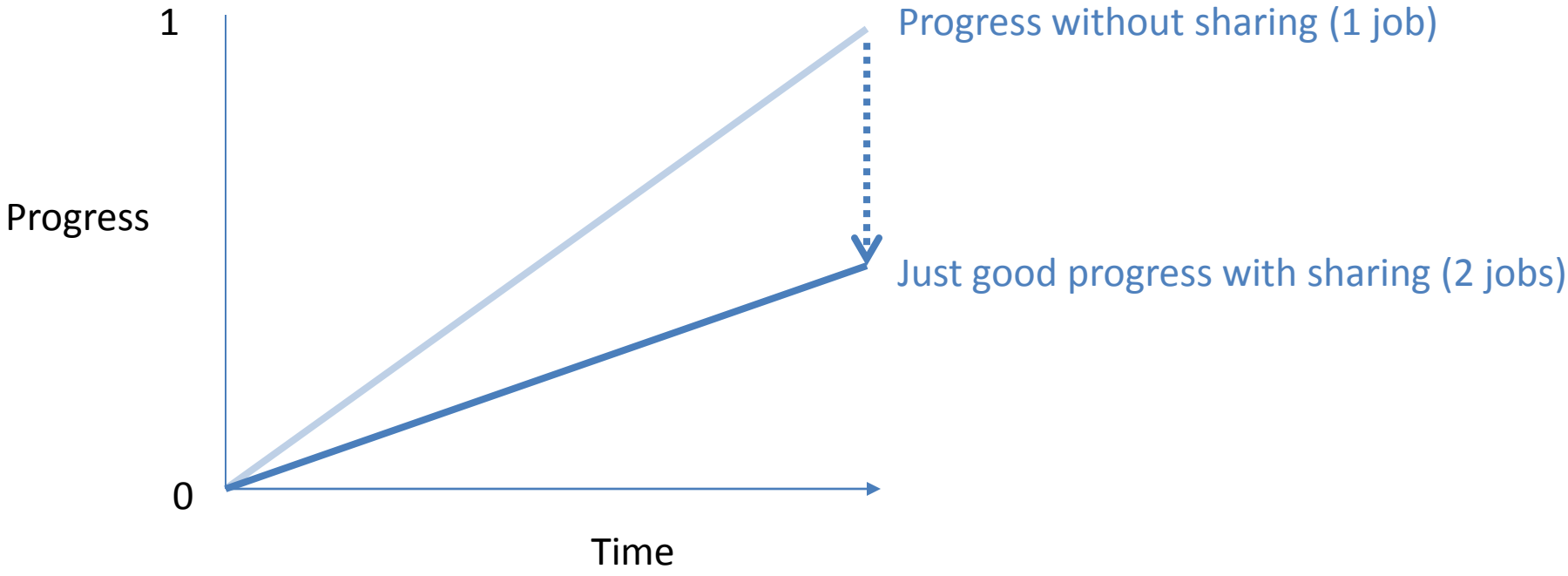
Fair Share Metric

- The scheduler try to equalize “share” of all jobs
  - SlotShare : Number of slots owned
    - Does not work well in heterogeneous environments
  - ProgressShare: Progress being made with owned slots / all slots
    - Contribution of a slot to a job’s progress rate

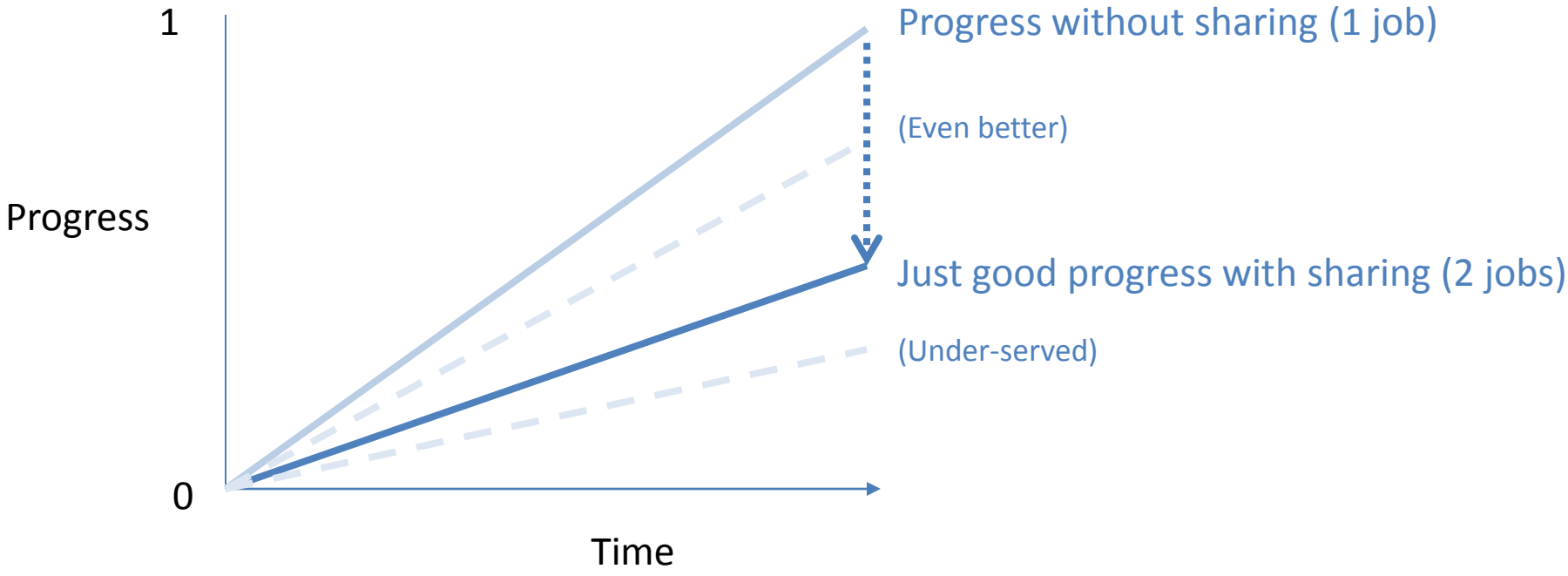
Progress Share



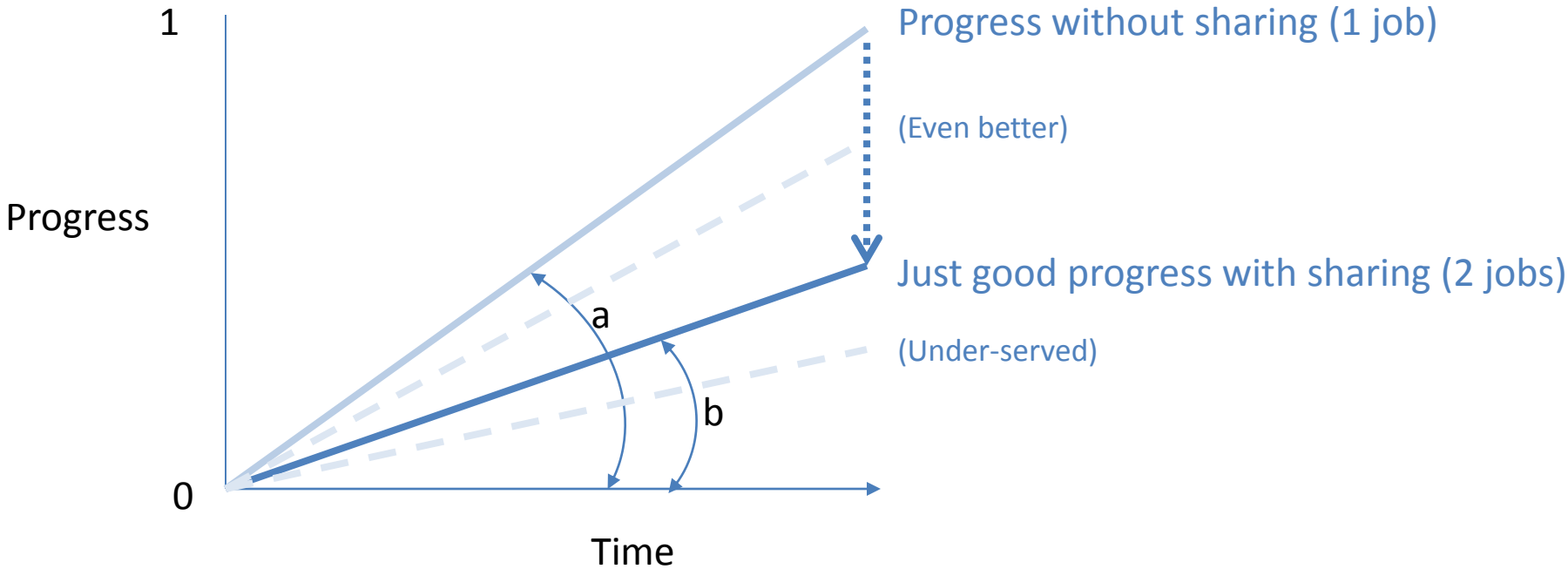
Progress Share



Progress Share

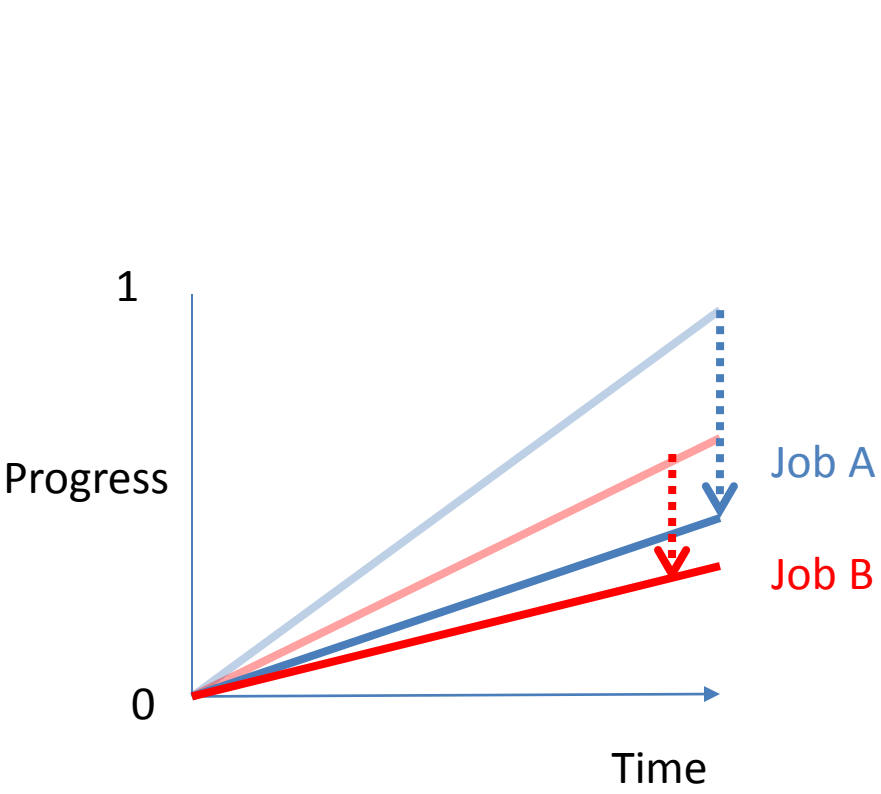


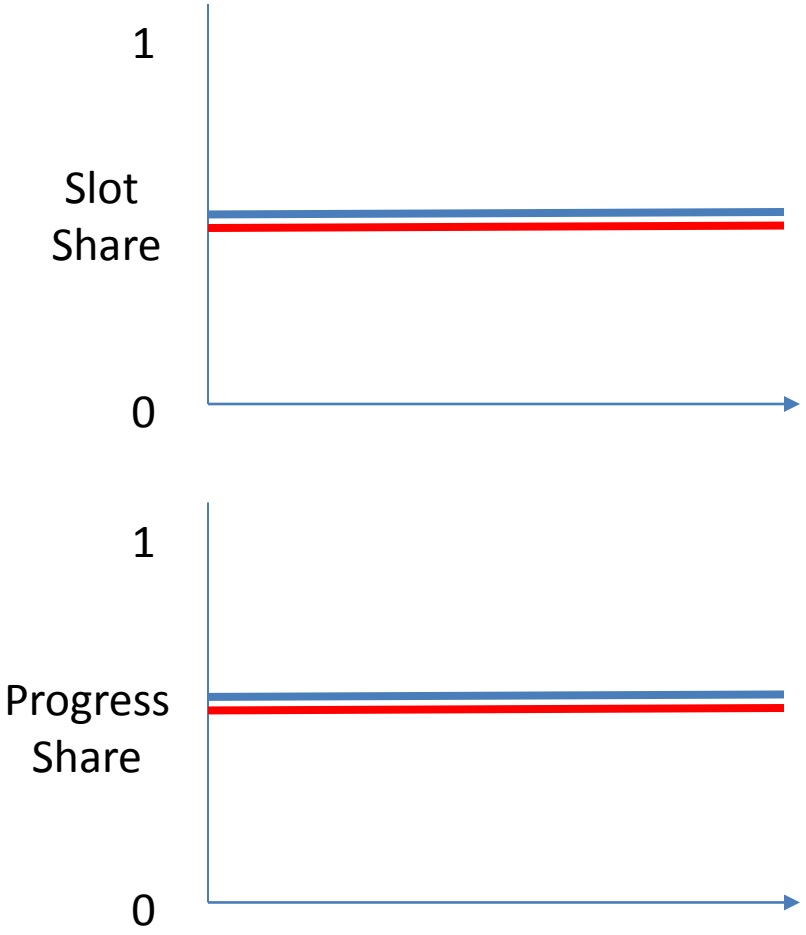
Progress Share



Progress Share of Job A =
Ratio of progress slope (b/a)

Homogeneous case



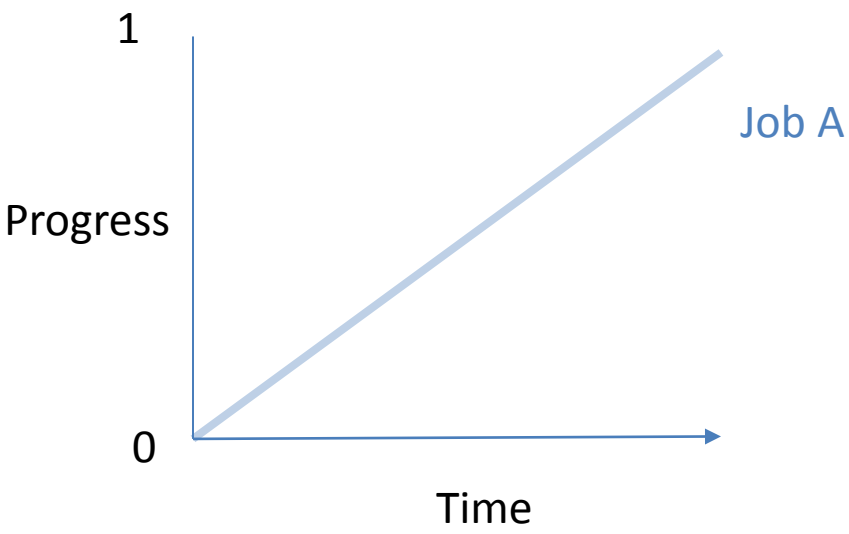


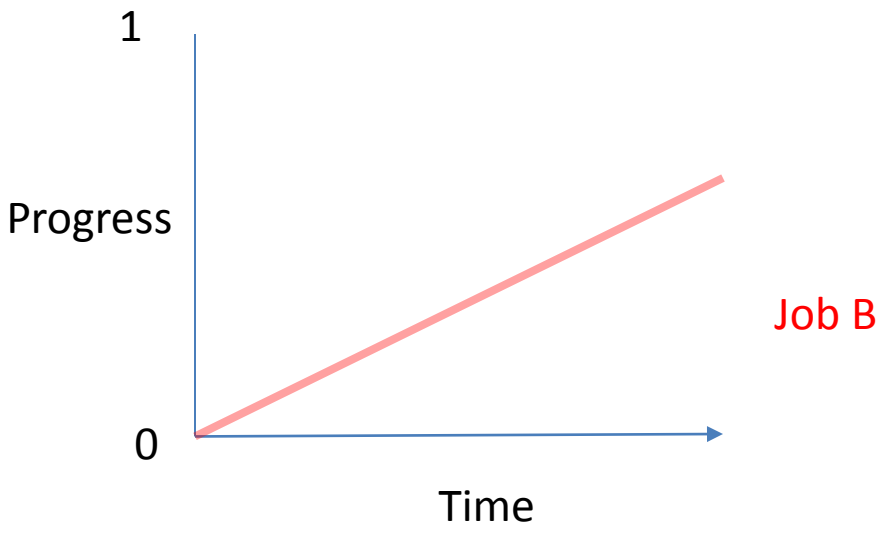
Heterogeneous case

Job A runs faster on gray slots

A			A	
A			A	
A			A	
A			A	
A	A	A	A	A
A	A	A	A	A

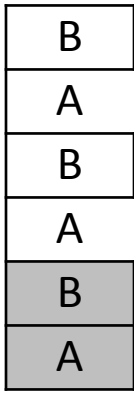
B	B	B	B	B
B	B	B	B	B
B	B	B	B	B
B	B	B	B	B
B	B	B	B	B
B	B	B	B	B

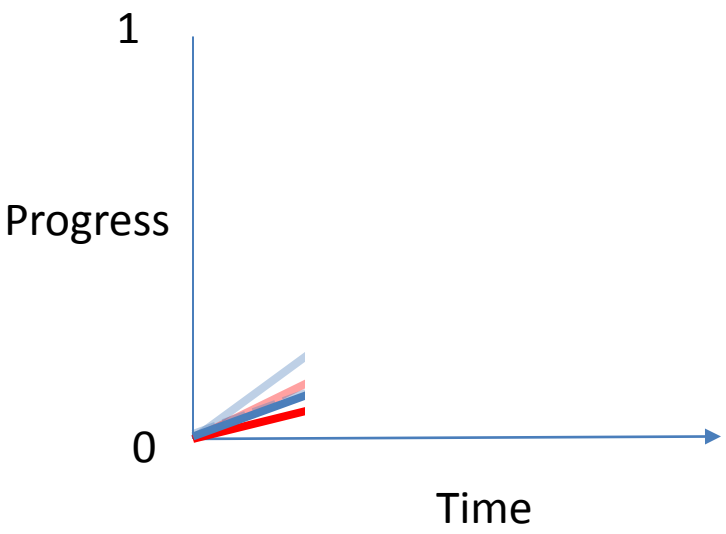


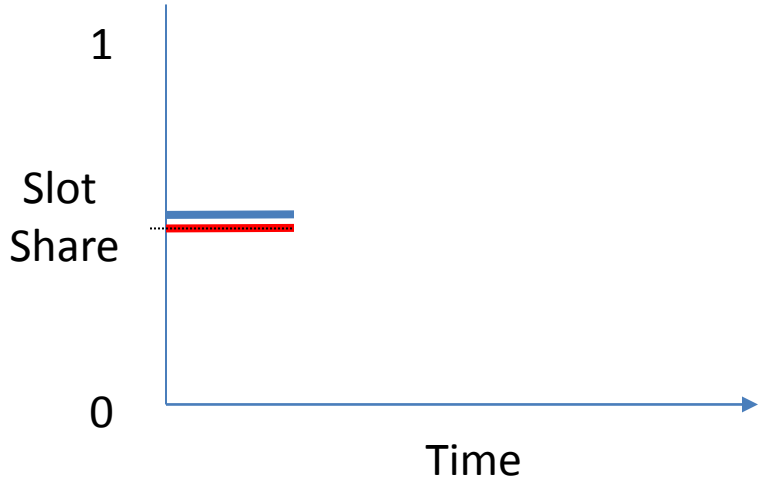


Heterogeneous case 1

Using SlotShare

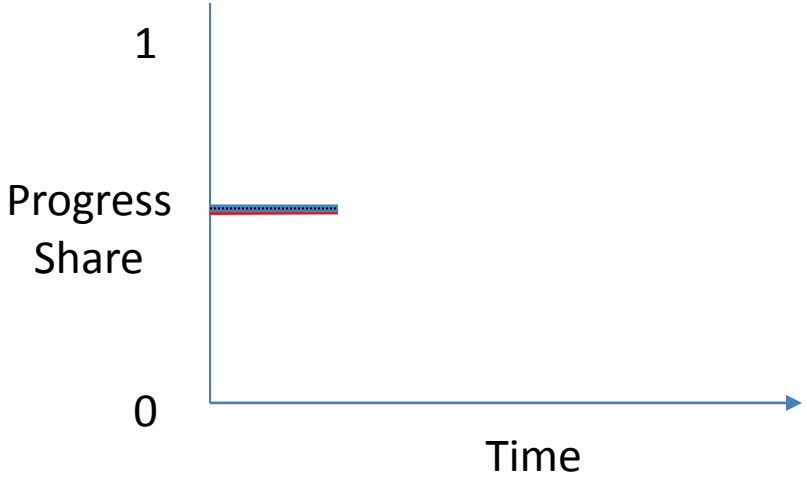






Job A

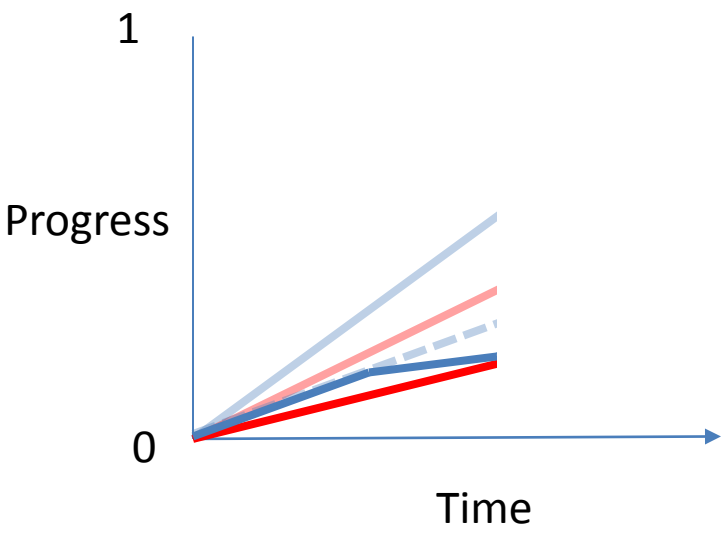
Job B



Heterogeneous case 1

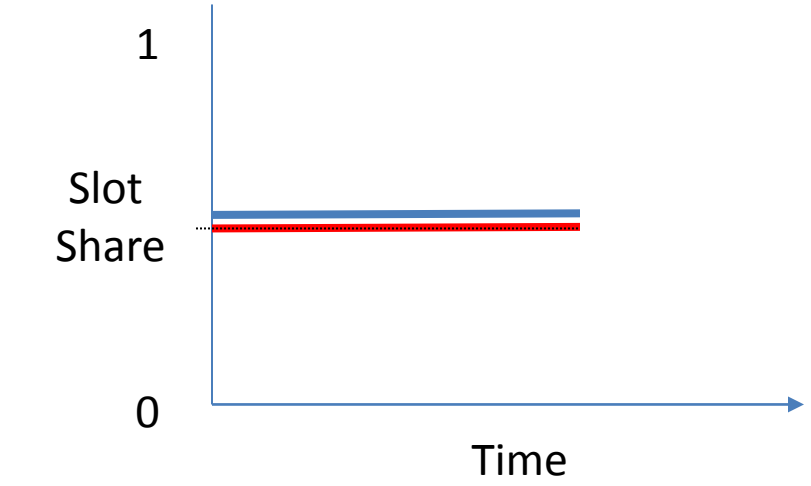
Using SlotShare

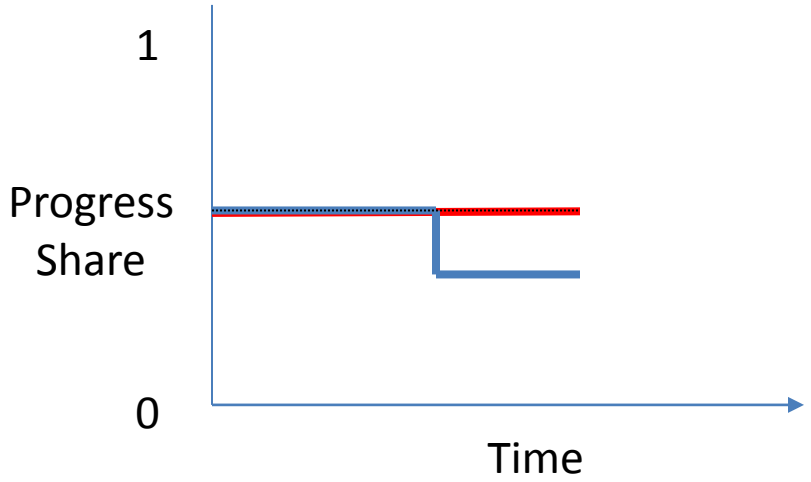
B	B	A
A		
B	B	B
A		
B	A	B
A	B	B



Job A

Job B



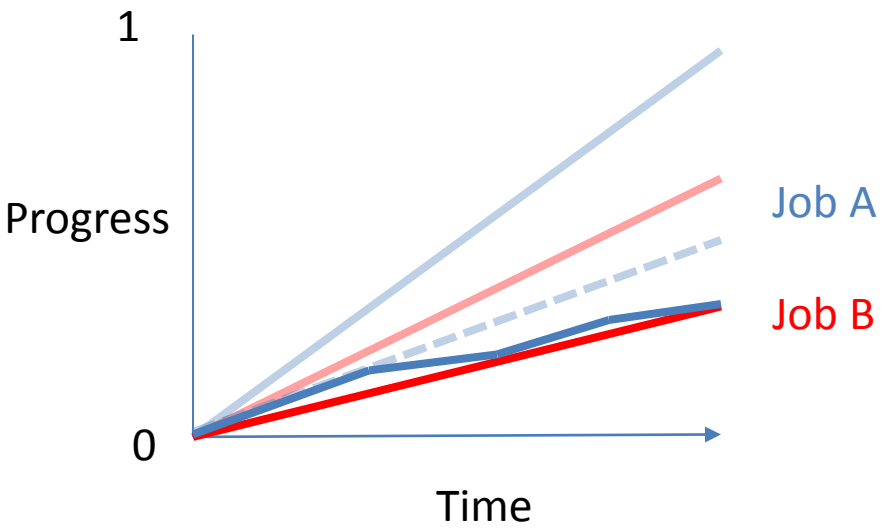


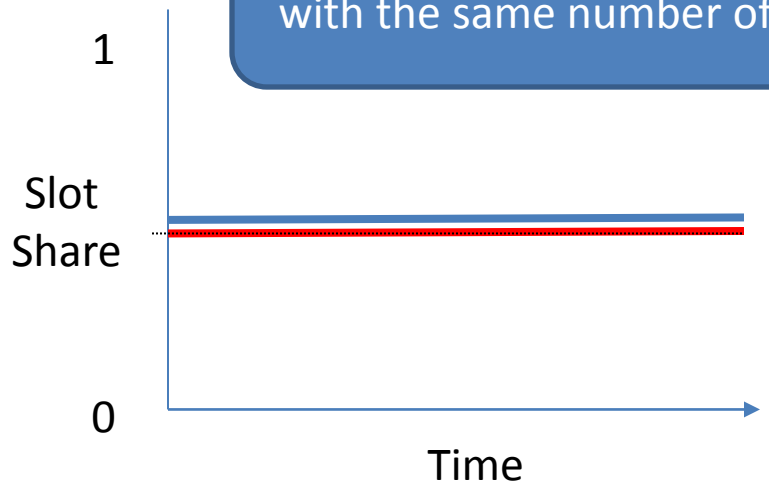
Heterogeneous case 1

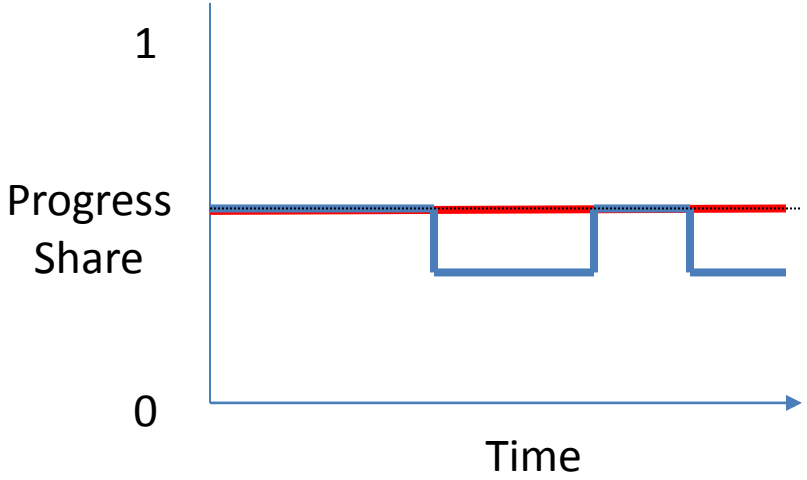
Using SlotShare

B	B	A		
A			B	B
B	B	B	A	
A			B	A
B	A	B	A	B
A	B	B	B	B

Job A is making less progress, with the same number of slots

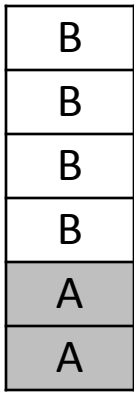


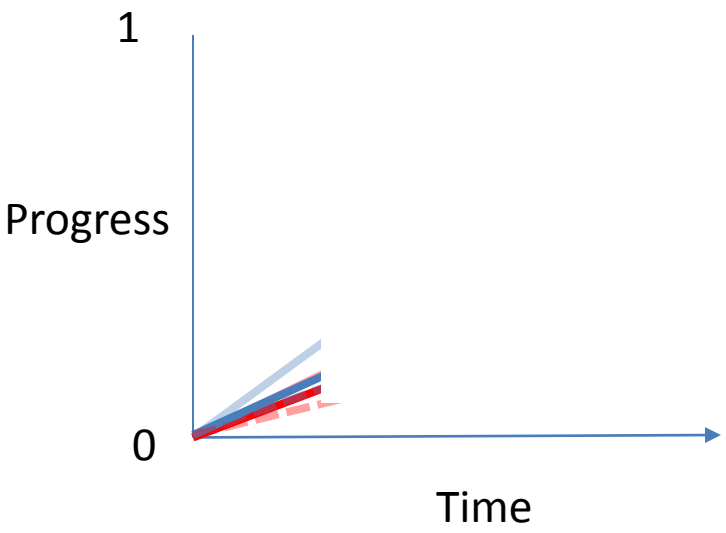




Heterogeneous case 2

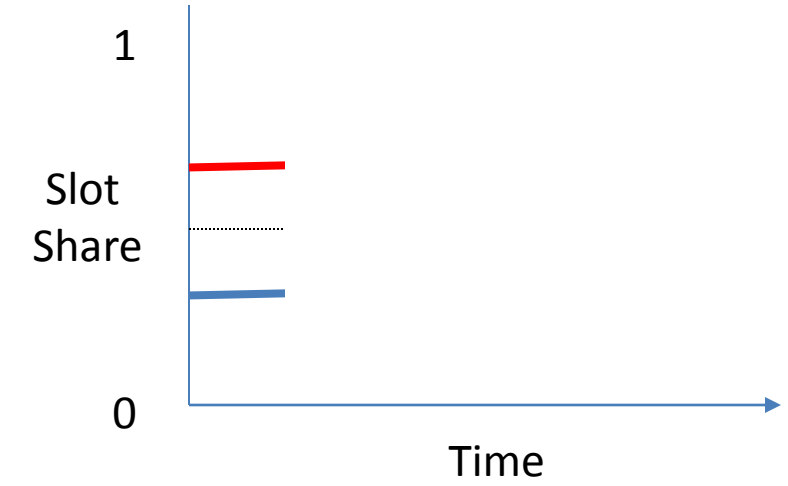
Using ProgressShare

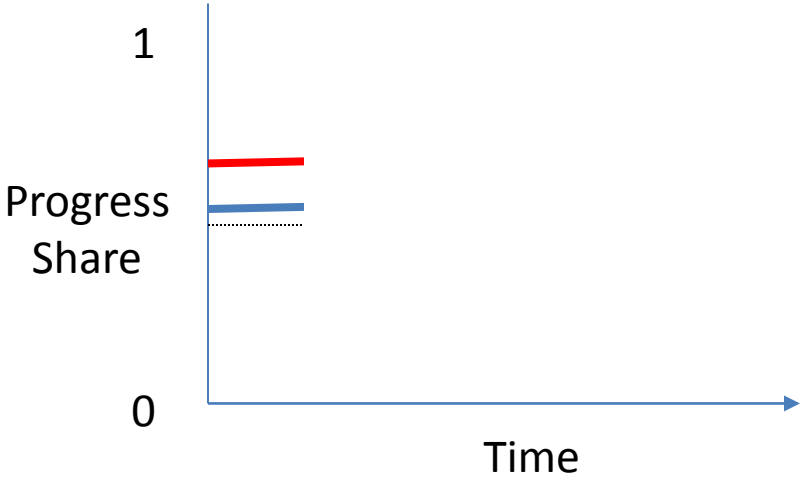




Job A

Job B

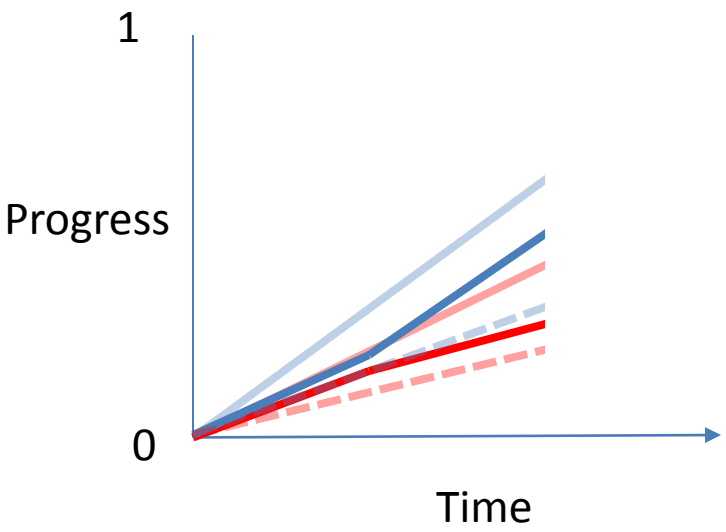




Heterogeneous case 2

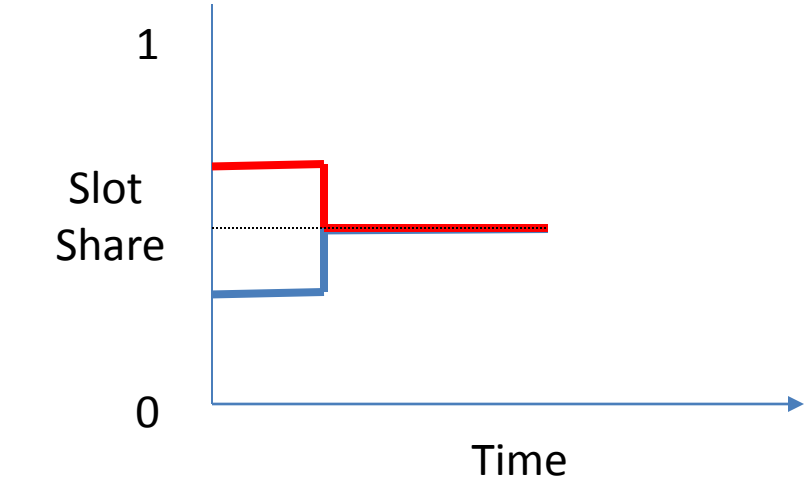
Using ProgressShare

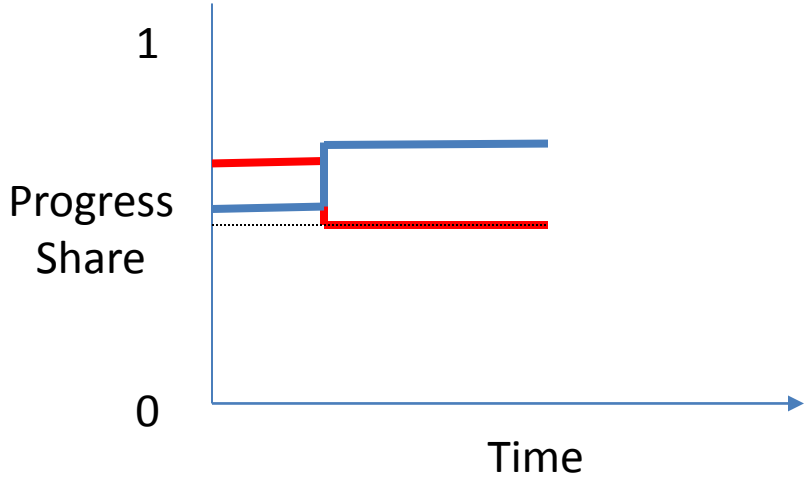
B	B	B
B	B	B
B	B	B
B	A	
A	A	A
A	A	A



Job A

Job B

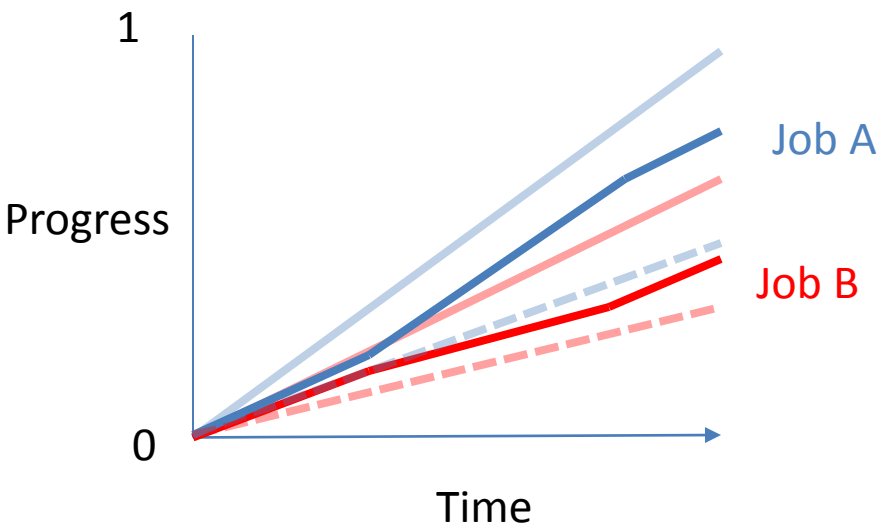




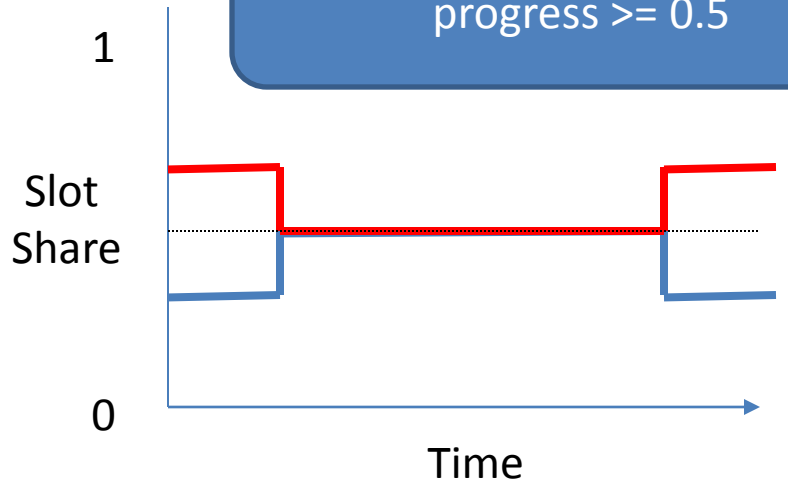
Heterogeneous case 2

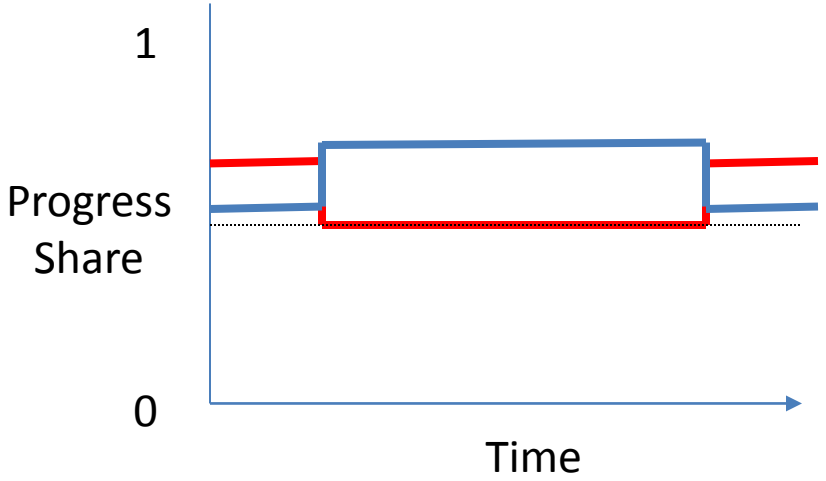
Using ProgressShare

B	B	B	B	B
B	B	B	B	B
B	B	B	B	B
B	A			B
A	A	A	A	A
A	A	A	A	A



Both jobs making progress >= 0.5





Performance Gain of Using Progress Share

Summary

- Heterogeneity should be taken account at both level of two-level scheduling
  - Resource Allocation and Job Scheduling
- Need to redefine “share” to provide performance and fairness simultaneously in heterogeneous environments
  - Propose “progress share”
- Future Work
  - Combine with sub-linear performance model
  - Consider inference of co-located jobs