

Elasticity in HCS

- Administrators can configure each queue with limits
 - Soft limit: how much % of cluster is the queue guaranteed to occupy
 - (Optional) Hard limit: max % of cluster given to the queue
- Elasticity
 - A queue allowed to occupy more of cluster if resources free
 - But if other queues below their capacity limit, now get full, need to give these other queues resources
- Pre-emption not allowed!
 - Cannot stop a task part-way through
 - When reducing % cluster to a queue, wait until some tasks of that queue have finished