Scheduling BoF
Problems

• Accelerating batch completion
• Adaptive replication wrt errors
• Making apps that use random seed reproducible
• Long-running jobs
Batch Completion

• Problem

• Delays due to job distribution

• Delays due to job error/failures
Proposed Solutions

- Adapt depending during lifecycle of batch of jobs
- More replication (angers volunteers)
- More predictable, faster hosts
- Dynamically divide work units

Client strategies

- Flag and increase priority of job on client
- Workunit score to indicate potential to be delayed on client (factors: “new host”, last connect time, availability logs, fraction done, current workload, surpassing of est_runtime)
Adaptive replication

- Questions about how it works wrt hosts with errors (documentation)
Long-runnings jobs

- Job takes 1 month to complete
- Want it to run continuously via migration
- Solution
  - message_from_host file to send the state file, indicating how far along it is in the computation
  - have optional download file to use this state file