Schedulers in Practice (contd.)

Low latency Scheduling

Announcements / Reminders

- HW was due before class.

- Proposals, unless you have gotten an extension, are due at midnight

- It's an excellent time to begin working on the project
  
  -> Best case is done early. More time to enjoy the warm weather.
  
  -> Likely case: spread out work so the end of semester is more fun.

Last Week

Update state of (current) week looks like
• What a cluster (approximately) looks like.
• Interactive vs Batch jobs
• Why schedulers need to be aware of job semantics
  → What this looks like for Dataflow jobs
• Two level schedulers.

Overview

Subtask

What Alice's tasks get?

Alice School
Beth School

Resource Allocation: What Alice/Beth gets?

What resources are available?

Isolation

• Remember our discussion about fairness.
  → Sharing incentive

• Assumption

No
Better
Off


Reality

Mechanisms

- Processes
- VMs
- Core affinity
- CFS/Scheduling policy
- New isolation hardware (SR-iov, MIG, CAT, ...)
- Queue management

Focus for today

Alice cannot affect Beth's utility if

- Connectedness
- Latency
- Throughput
- Security

Both ran on

½ cluster
Quick Recap: What is a Process?

- PCB
  - File Table
  - Page Table
- TCB
  - Context
- TCB
  - Context
- TCB
  - Context
  - CPU registers

Brief Aside: Difference from VMs

Guest Kernel

PCB (or P)

VM
Quick Recap

Direct cost when switching from one process to another.

 Aside: Container

PCB for VM → TCB → Context → VM PT → FP
Implications

1. For tasks running on CPUs?
   - Cost to sharing

2. For new hardware?

Indirect costs
Why Look At These Mechanisms: Latency Sensitive Jobs

Why Rel. JCT?

$$|F_0| = |F_1| = 0.5 |S_0| = T$$

Analysis:

Optimize

Relo Job Completion Time (JCT)

= Time Waiting + Execution Time

S_0 - \frac{2T}{T+1} = 1

F_0 - \frac{2T+1}{T+1} = 3

F_1 - (2T+1-T+1)
Complicating Things Further

- Don't usually know task lengths a-priori
- Have multiple cores

Preemption Based Approaches
Request Segregation - Persephone?

How far can we push this?