Advanced disk scheduling

"Freeblock scheduling"

Eno Thereska

(slide contributions by Chris Lumb and Brandon Salmon)

PARALLEL DATA LABORATORY

Carnegie Mellon University

Outline

- Freeblock scheduling: some theory
- Freeblock scheduling: applied
- Some details
- Q & A

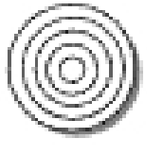
Some theory: preview

- Next few slides will review & show that:
  - disks are slow
    - mechanical delays (seek + rotational latencies)
  - there is nothing we can do during a seek
  - **there is a lot we can do during a rotation**
    - rotational latencies are very large
    - while rotation is happening go to nearby tracks and do useful work
  - "*freeblock scheduling*" = utilization of rotational latency gaps (+ any idle time)

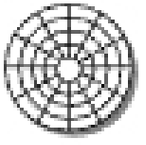
Are disks slow?

- Are the xfer speeds that slow?
  - no, xfer speeds of 200MB/s are pretty good
- So what is slow?
  - workload often not sequential
  - disk head has to move from place to place
  - seek (~ 4ms) + rotation (~ 3ms)
- Effective bandwidth can be very low
  - ~ 10-30MB/s
  - even when SPTF is used

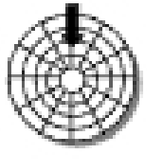
Surface organized into tracks



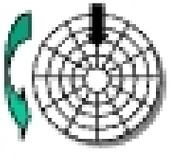
Tracks broken up into sectors



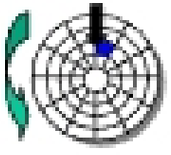
Disk head position



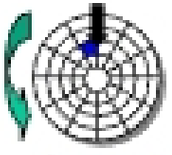
Rotation is counter-clockwise



About to read blue sector

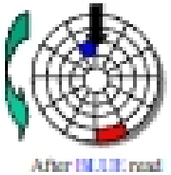


After reading blue sector



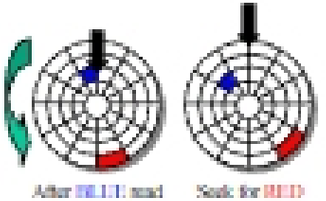
After BLUE read

Red request scheduled next



After BLUE read

Seek to Red's track



After BLUE read Seek for RED