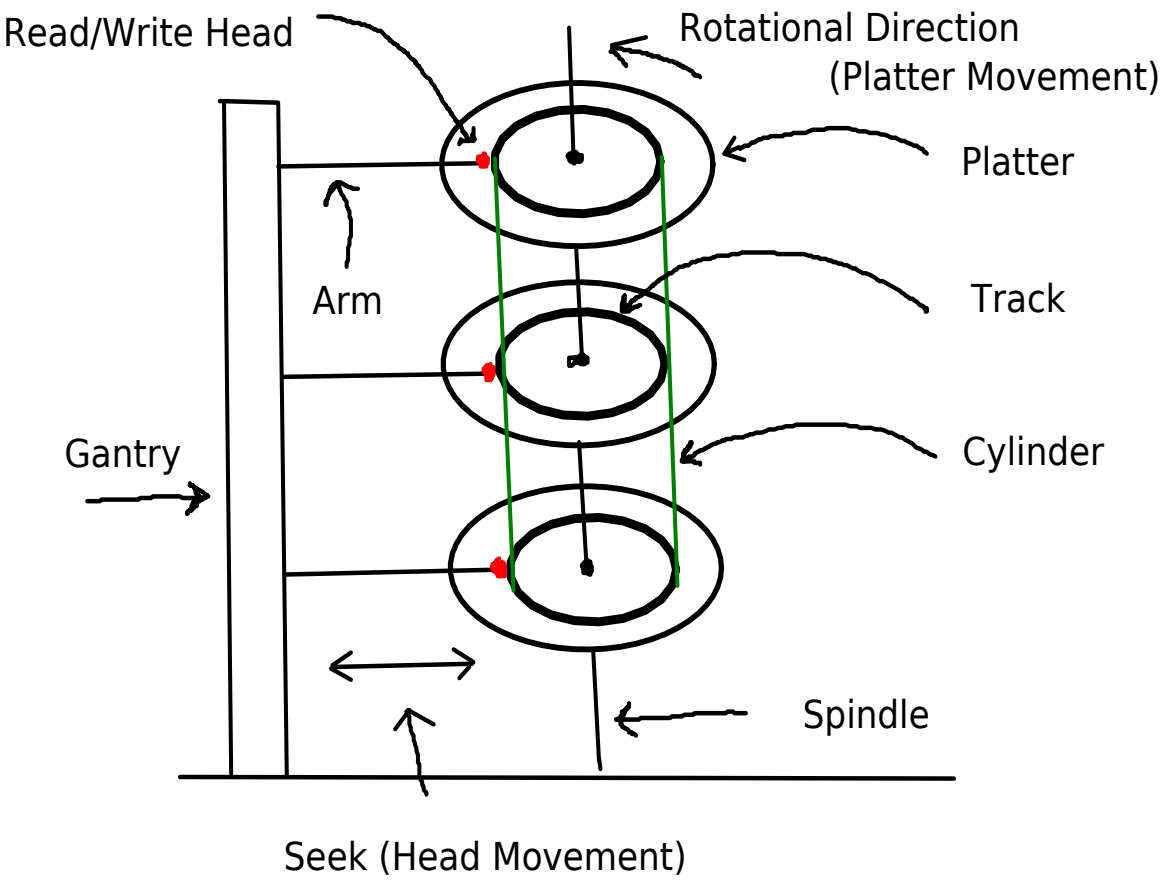
# Computer Science CSCI 251

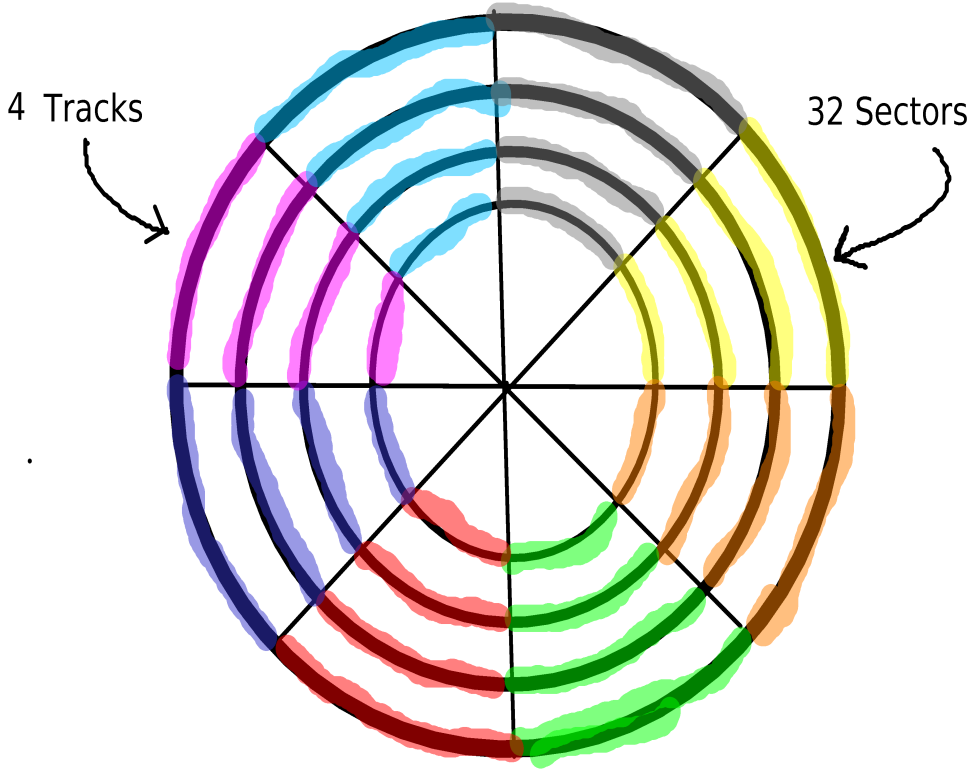
# Systems and Networks

*Dr. Peter Walsh*
*Department of Computer Science*
*Vancouver Island University*
*peter.walsh@viu.ca*

# Disk Architecture



# Disk Architecture cont.



# Terminology

- drives have 1 or more platters
- each platter has 1 or 2 sides
- each side has a read/write head
- each side is divided into concentric tracks
- platers and read/write heads are stacked
- a cylinder is the set of tracks the can be accessed simultaneously
- each track is divided into sectors
- moving a head over a track (or cylinder) is called seeking
- moving a head over a sector typically requires seeking and platter rotation

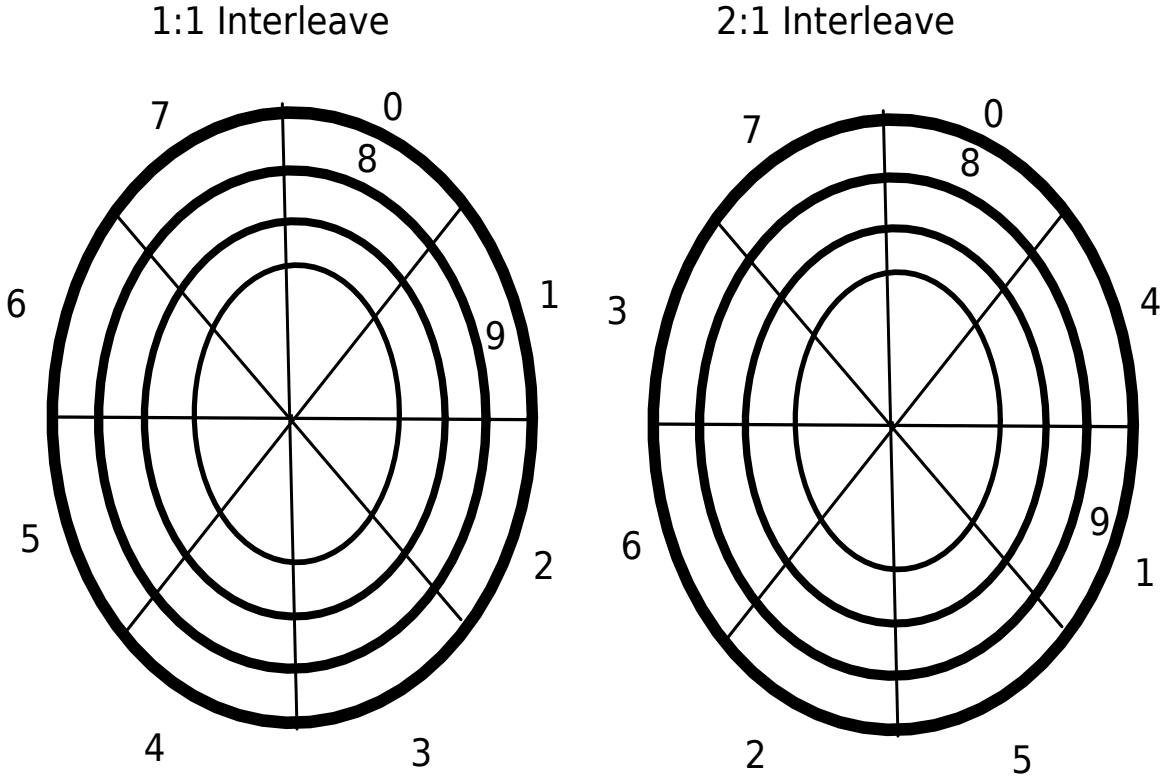
# Sectors

- Size
  - smallest addressable part of the disk for read/write
    e.g., 512 bytes
  - Unix uses the term block (1 or more sectors)
- Address
  - `cylinder#, head#, sector#`
- Access Time
  - seek: move head to track/cylinder
  - rotational delay (latency): spin sector
    under the head
  - transfer: time to read/write data
- Old Drives
  - rotational delay (latency) << seek time
- Modern Drives
  - rotational delay (latency) $\approx$ seek time

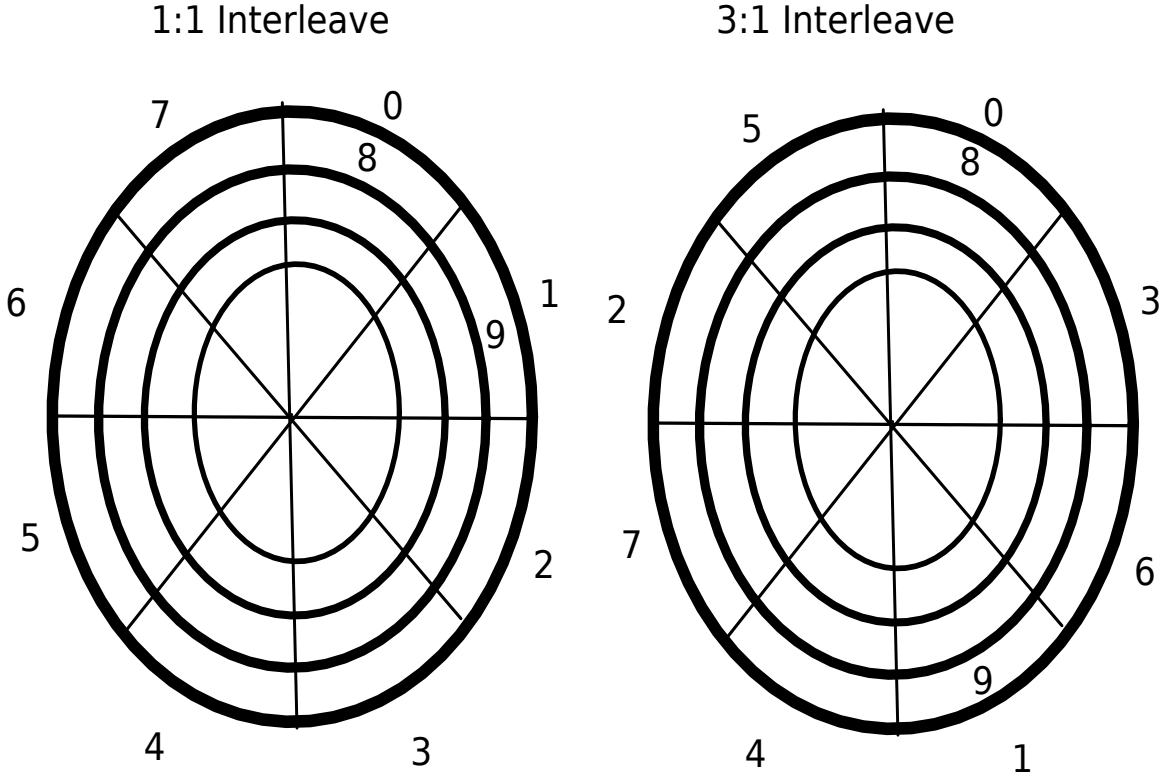
# Disk Architecture Optimizations

- Sector Interleaving
  - because of the overhead associated with read/write transfer, it may be more efficient to interleave sectors that are read in sequence
- Cylinder Skew
  - is a form of interleaving employed to decrease overhead when accessing consecutive sectors that cross track boundaries
- Multi Zoned
  - outer tracks are bigger and may be configured to contain more sectors that smaller inner tracks

# Sector Interleaving

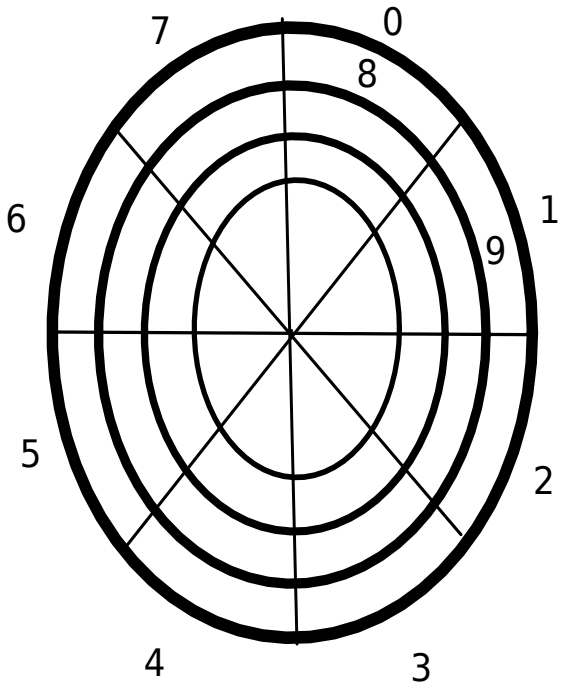


# Sector Interleaving cont.

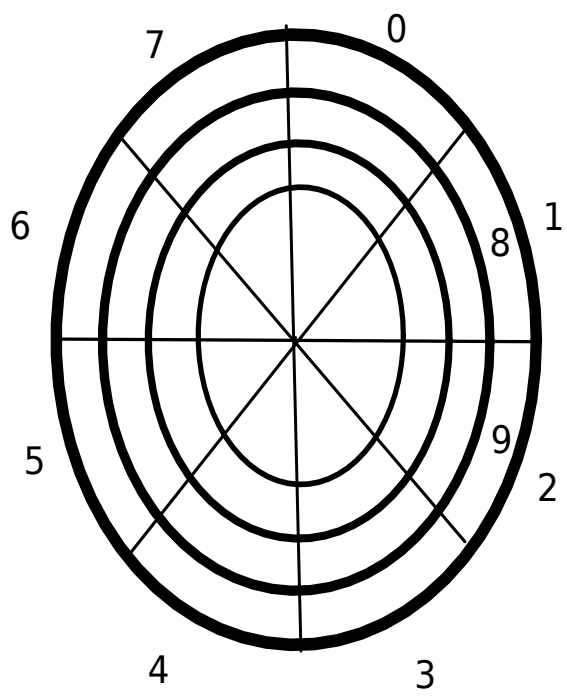


# Cylinder Skew

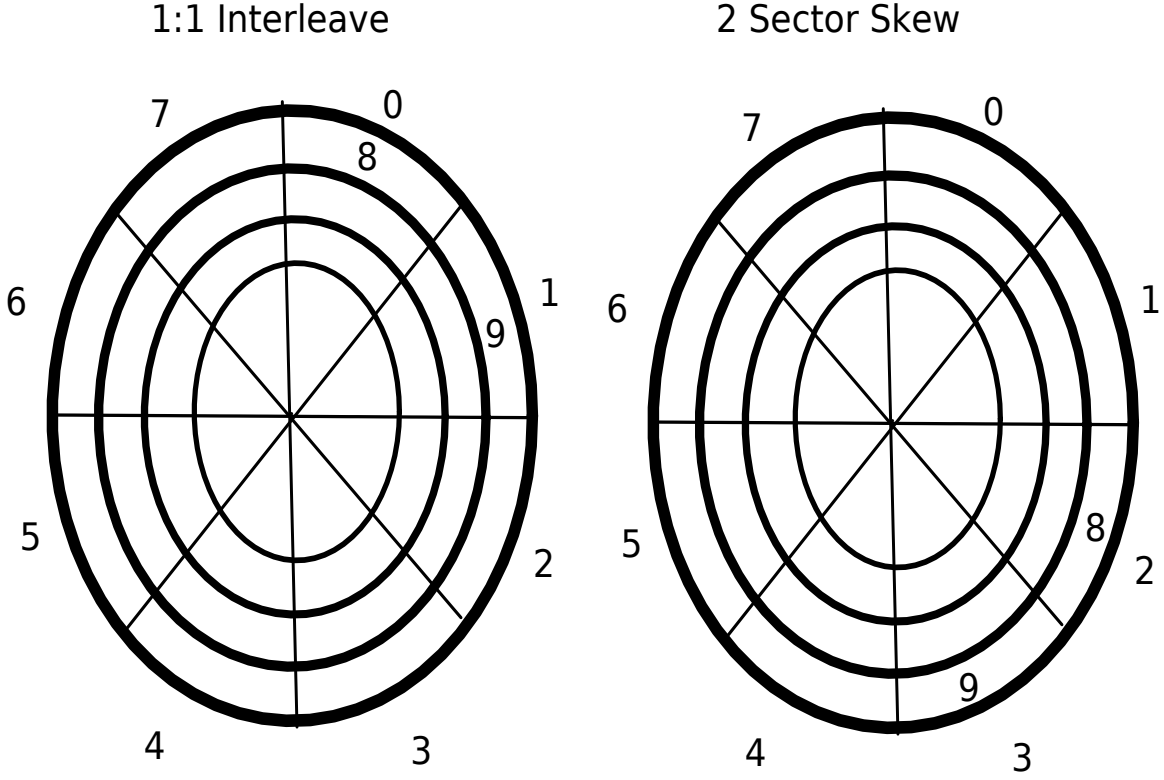
1:1 Interleave



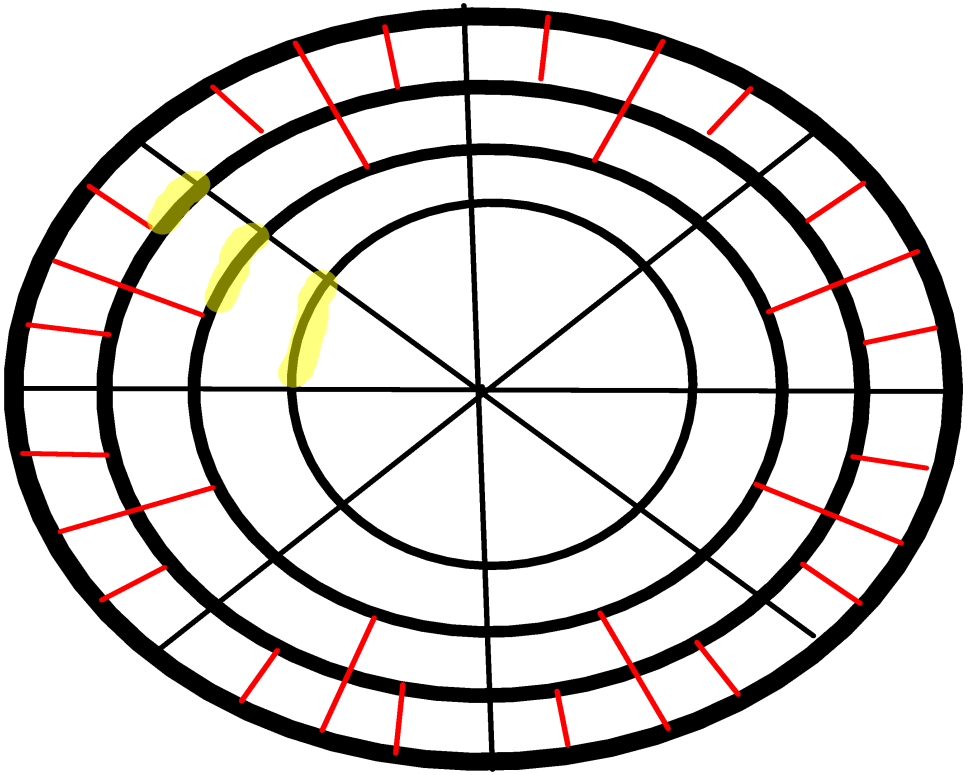

1 Sector Skew



# Cylinder Skew cont.



# Multi Zoned



# Workloads

- Random
  - issue reads to random locations on disk
  - e.g., database management system (DBMS)
- Sequential
  - read a large number of sectors consecutively
  - e.g., sequential file processing

# I/O Time

| (Fig. 37.5 pp 463) | Cheetah | Barracuda |
|---|---|---|
| RPM | 15,000 | 7,200 |
| Average Seek | 4ms | 9ms |
| Max Transfer | 125 MB/s | 105 MB/s |

Random Workload - 4 KB Reads
Sequential Workload - 100 MB Read

$$T_{IO} = T_{seek} + T_{rotation} + T_{transfer}$$

$$R_{IO} = \frac{S_{transfer}}{T_{IO}}$$

# Random Workload (Cheetah)

$T_{seek} = 4\,ms$

RPM = 15,000

$\frac{1}{2}$ rotation $= \frac{60}{30,000}\,s = 0.002\,s$

$\Rightarrow T_{rotation} = 2\,ms$

Max Rate = 125 MB/s

$T_{transfer}(4KB) = \frac{4,000}{125,000,000} = 32\,\mu s$

# Random Workload (Cheetah)

$$T_{IO} = T_{seek} + T_{rotation} + T_{transfer}$$

$$\approx 6 \text{ ms}$$

$$R_{IO} = \frac{S_{transfer}}{T_{IO}}$$

$$= \frac{4}{6} \text{ KB/ms} = 0.66 \text{ MB/s}$$

# Random Workload (Barracuda)

$T_{seek} = 9\ ms$

RPM = 7,200

$\frac{1}{2}$ rotation $= \frac{60}{14,400}\ s = 0.0041\ s$

$\Rightarrow T_{rotation} = 4.1\ ms$

Max Rate = 105 MB/s

$T_{transfer}(4KB) = \frac{4,000}{105,000,000} = 38\ \mu s$

# Random Workload (Barracuda)

$$T_{IO} = T_{seek} + T_{rotation} + T_{transfer}$$

$$\approx 12 \text{ ms}$$

$$R_{IO} = \frac{S_{transfer}}{T_{IO}}$$

$$= \frac{4}{12} \text{ KB/ms} = 0.33 \text{ MB/s}$$

# Sequential Workload (Cheetah)

$T_{seek} = 4\,ms$

RPM = 15,000

$\frac{1}{2}$ rotation $= \frac{60}{30{,}000}\,s = 0.002\,s$

$\Rightarrow T_{rotation} = 2\,ms$

Max Rate = 125 MB/s

$T_{transfer}(100MB) = \frac{100}{125}\,s = 800ms$

# Sequential Workload (Cheetah)

$$T_{IO} = T_{seek} + T_{rotation} + T_{transfer}$$

$$\approx 800 \text{ ms}$$

$$R_{IO} = \frac{S_{transfer}}{T_{IO}}$$

$$= \frac{100}{800} \text{ MB/ms} = 125 \text{ MB/s}$$

# Sequential Workload (Barracuda)

$T_{seek} = 9\ ms$

RPM = 7,200

$\frac{1}{2}$ rotation $= \frac{60}{14{,}400}\ s = 0.00417\ s$

$\Rightarrow T_{rotation} = 4.1\ ms$

Max Rate = 105 MB/s

$T_{transfer}(100\ MB) = \frac{100}{105}\ s = 952\ ms$

# Sequential Workload (Barracuda)

$$T_{IO} = T_{seek} + T_{rotation} + T_{transfer}$$

$$\approx 950\,ms$$

$$R_{IO} = \frac{S_{transfer}}{T_{IO}}$$

$$= \frac{100}{950}\,MB/ms \approx 105\,MB/s$$

# Comparison Summary

(Fig. 37.6 pp 464)

| | Cheetah | Barracuda |
|---|---|---|
| RPM | 15,000 | 7,200 |
| Average Seek | 4ms | 9ms |
| Max Transfer | 125 MB/s | 105MB/s |
| Random Workload - 4 KB Reads | | |
| Sequential Workload - 100 MB Read | | |
| I/O Rate (Random) | 0.66 MB/s | 0.31 MB/s |
| I/O Rate (Sequential) | 125 MB/s | 105 MB/s |

# Disk Scheduling

- Old Drives
  - rotational delay (latency) << seek time
  - OS responsible for disk scheduling
- Modern Drives
  - rotational delay (latency) $\approx$ seek time
  - disk drive controller more involved in scheduling

# Disk Scheduling cont.

- Shortest Seek Time First (Scan)
  - sweep the disk servicing requests in-order across the tracks
- Shortest Position Time First
  - needs disk controller support

- OS/Controller Cooperation
  - OS selects a set of I/O requests to be serviced optimization is based on OS criteria
  - disk controller schedules requests for service optimization based on detailed knowledge of disk drive architecture and performance