

1. Last time

2. Disks (HDDs)

Next time: intro to file systems

2. Disks

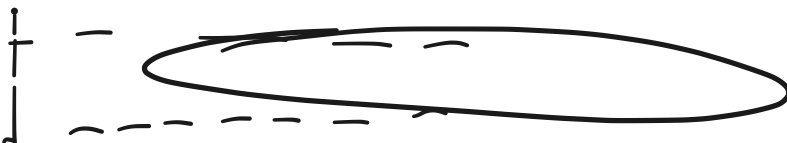
- 1. A. What is a disk?
- 1. B. Geometry
- 1. C. Performance ←
- 1. D. Common numbers
- 1. E. Interface to disk
- 1. F. Performance II
- 1. G. Disk scheduling (performance III)
- 1. H. Technology + systems trends

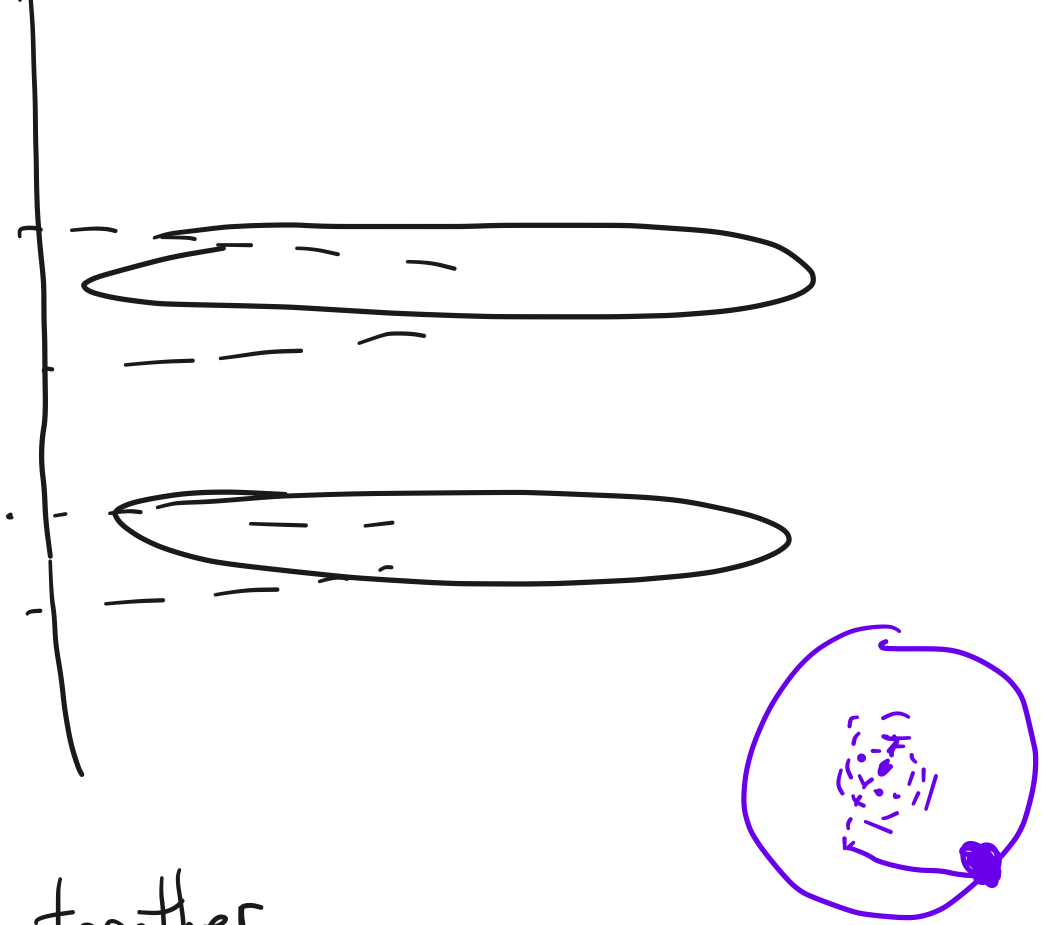
\$/GB

- stack of magnetic platters

- rotate together on central spindle

3600 - 15,000 RPM (=60 - 250 rot/sec)





Arms move together

Arms contain disk heads; heads read + write to platters

Geometry

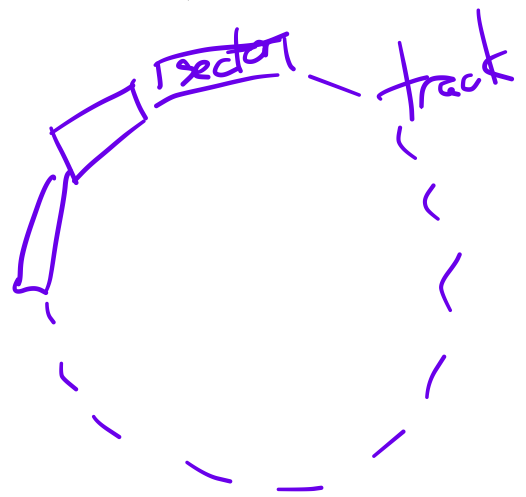
- track: circle on a platter

- sector: chunk of a track

- cylinder: all tracks of fixed radius on all platters

- generally only one head active at a time

- disk positioning system: moves head to a track and keeps it there



· seek: 4 phases: speedup, coast, slowdown, settle

Performance

Components of total transfer time:

rotational delay

seek delay

transfer time

"Avg seek time"

"time to seek $\frac{1}{3}$ of the disk", but
manufacturers might report:

$\frac{1}{3}$ of "time to seek the whole disk"



Common #'s

Capacity: TBs common (10^{12} bytes vs. 2^{40} bytes)

Platters: 8

of thousands or more

Number of cylinders: tens of thousands

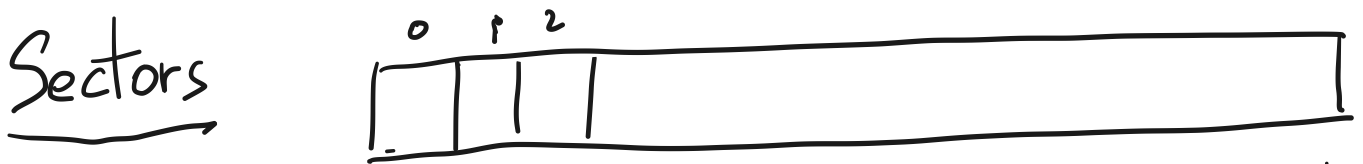
Sectors per track: ~ 1000

RPM: 10,000

Transfer rate: 50 - 150 MB/s

MTBF: ~ 1 million hours

How driver interfaces to disk

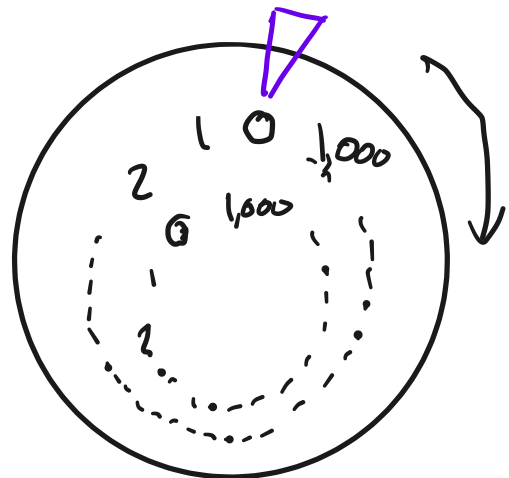


Interface to disk is a linear array of sectors

Sector: 512 bytes, moving to 4KB written atomically

Disk does some cool things under the hood (invisibly to OS)

- Zoning
- Skewing
- Sparring



Disk performance example

Spindle speed: 12,000 RPM

Avg. seek time: 12ms

Transfer rate: 128 MB/s

Sector: 512 bytes

(a) What is the throughput if doing 500 sector reads, spread randomly over the disk and serviced in FIFO order?

(b) Same question, but now the reads are sequential

(a)

$$\text{TPUT} = \frac{\text{bytes}}{\text{time}} = \frac{500 \times 512 \text{ bytes}}{500 \text{ random reads}}$$

1 read = rot. delay + seek time + transfer time

$$1 \text{ rot.} = \frac{60 \text{ s}}{\text{min}} \times \frac{\text{min}}{12,000 \text{ rot}} = \frac{6}{1200} \approx 5 \text{ ms}$$

$$\text{avg. rot. delay} = 2.5 \text{ ms}$$

$$\text{avg. seek time} = 12 \text{ ms}$$

$$\text{transfer time} = \frac{1 \text{ sec}}{128 \text{ MB}} \times 512 \text{ bytes} = \frac{2^9}{2^{27}} = 2^{-18} \approx 4 \mu\text{s}$$

$$\frac{500 \times 512 \text{ bytes}}{500 \times (2.5 \text{ ms} + 12 \text{ ms} + 4 \mu\text{s})} \approx \frac{512 \text{ bytes}}{14.5 \text{ ms}} \approx \frac{512 \times 70 \text{ bytes}}{\text{s}}$$

$$\left\{ \frac{1}{14.5 \text{ ms}} \approx \frac{70}{70 \times 14.5 \text{ ms}} = \frac{70}{1000 \text{ ms}} = \frac{70}{\text{s}} \right\}$$

$$\approx 500 \times 70 \frac{\text{bytes}}{\text{sec}} = \frac{35000 \text{ bytes}}{\text{sec}} = \boxed{\frac{35 \text{ KB}}{\text{sec}}}$$

(b)

$$\text{TPUT} = \frac{\text{bytes}}{\text{time}} = \frac{500 \times 512 \text{ bytes}}{1 \text{ seq. read}}$$

$$\text{avg. rot. delay} = 2.5 \text{ ms}$$

$$\text{avg. seek time} = 12 \text{ ms}$$

$$\text{transfer time} = \frac{1 \text{ sec}}{128 \text{ MB}} \times 500 \times 512 \text{ bytes} \approx \frac{2^9 \times 2^9}{2^{27}} = 2^{-9}$$

$$= 2 \text{ ms}$$

$$\begin{aligned}
 & \left. \begin{array}{l} \text{ } \\ \text{ } \\ \text{ } \end{array} \right\} \rightarrow \frac{500 \times 512 \text{ bytes}}{2.5 \text{ ms} + 12 \text{ ms} + 2 \text{ ms}} = \frac{500 \times 512 \text{ bytes}}{16.5 \text{ ms}} \\
 & \approx \frac{30 \times 512 \text{ bytes}}{1 \text{ ms}} \\
 & = \frac{3 \times 5 \text{ KB}}{1 \text{ ms}} = \frac{15 \text{ MB}}{\text{s}}
 \end{aligned}$$

~~X~~

 500 (_ + _ + _)

~~X~~

 _ + _ + 500 · _

