Today

⇒ I/O
    input/output

⇒ devices

⇒ case studies:
    hard drive
    flash-based SSD
Device: API, internals

interface

generic device interface:

command register

data register

status register

firmware

internal: memory, control logic (CPU)
Device interaction protocol:

1. **check status** → data
   - \( \text{while} (\text{status} == \text{Busy}) \)
   - \( ; \text{ // spin, poll (ask)} \)

2. **make a request**
   - \( \rightarrow \text{put (write) data} \)
   - \( \rightarrow \text{put command} \)
   - \( \rightarrow \text{command register} \)
   - \( \text{(now device goes to work)} \)

3. **wait for completion**
   - \( \text{while} (\text{status} == \text{Busy}) \)
   - \( ; \text{ // spin, poll)} \)

**Registers:**
- \( \rightarrow \text{status} \)
- \( \rightarrow \text{command} \)

**How to read/write device registers?**
- **Explicit I/O**
  - \( \bigcirc \text{ instructions} \)
  - \( \bigcirc \text{ (in, out)} \)
- **Memory-mapped I/O**
Problems w/ basic protocol:

- nesting time => polling
  (huge waste of CPU cycles)

interrupt:
- program device to interrupt the CPU when operation is complete

```
[while (status == Busy)
  ; 4 bytes @ time
  write data
  write cmd
[while (status == Busy)
  ;
  [billions/sec]

CPU
```

w/ polling

CPU

Device

done? no! no! still

w/w

... na na

no go again

yes

w/ interrupts

OS: switch to B

CPU A

(int)

Device

overlap CPU / I/O

(multiprogramming)
interrupt handling:

⇒ OS in charge

@ boot time!
  installs interrupt handlers
    (code that runs when
     int occurs)

@ run time:
  to issue an I/O ⇒
    system call ⇒ ask OS to
       perform I/O

OS: issue I/O,
   switch to some other
    program

OS: interrupt,
   may continue running
original program
Copy data to device: Slow

Time

Copy is slow: CPU has to move each byte (programmed I/O)
Solution:

DMA controller
( another device )

\[ \Rightarrow \] handles copying data
to/from device and main memory
between cpu (os) and memory

when it needs to transfer:
programs DMA device to do it

( start, length, device )

[Diagram of data transfer from device to memory]
Hard Disk Drives (HDDs)

Persist (non-volatile)

Persistent storage

Turn off power, info does not go away

(contrast: main memory)

=) Cheapest, reasonably performing kind of persistent storage
(data centers... )
How hard drive operates:

- Platter (many, or one) spins
- Coated magnetic surface
- Disk arm
- Head: sense bits (read), change (write) (above surface)
- Disk arm move:
  - Seek
  - Wait: rotation
- Transfer: do read/write
Spin @ fixed rate:

- 7200 Rotations/minute (RPM)
- 10000 RPM

One rotation: [6 ms]

Data laid out on surface:

- Platter
- Spindle
- Track
- Sector (512 bytes)

Minimal size: read/write

⇒ single sector
drive: can have multiple platters

Side view:

- r/w head
- platter
- motor
I/O:

Steps:

=> **Seek**: move to desired track
   (accelerate, coast, decelerate, settle)
   [slow: milliseconds]

=> rotational delay: wait

=> transfer: read/write
Two different workloads:

- read random 4KB (small) blocks
- read random 4MB (large) blocks

Disk:
- Average random seek: 7 ms
- Average rotation: 3 ms
- Transfer at rate: 100 MB/s
avg seek: 7 ms
avg rot: 3 ms

\[ T_{I/O} = T_{\text{seek}} + T_{\text{rot}} + T_{\text{transfer}} \]

\[ (4 \text{ KB}) = 7 \text{ ms} + 3 \text{ ms} + \approx 0 \text{ ms} \]

4 KB transfer @ 100 MB/s

\[ \Rightarrow 4 \text{ KB} \cdot \frac{1 \text{ MB}}{100 \text{ MB}} \cdot \frac{1 \text{ MB}}{1024 \text{ KB}} \cdot \frac{1000 \text{ ms}}{1 \text{ s}} \]

\[ \Rightarrow \frac{4}{100} \text{ ms} \Rightarrow \boxed{0} \]
4 MB I/O:

\[ 7 \text{ ms} + 3 \text{ ms} + \frac{40 \text{ ms}}{(\text{rot})} \] (seek)

\[ \frac{4}{100} \text{ ms} \] (4 KB) + transfer

\[ \approx \frac{4000}{100} \text{ ms} \] (4 MB) = \[40 \text{ ms}\]

4 MB transfer

4 MB I/O

4 KB I/O

Rate: \[ \frac{4 \text{ MB}}{50 \text{ ms}} \Rightarrow 80 \text{ MB/s} \]

Rate: \[ \frac{4 \text{ KB}}{10 \text{ ms}} \Rightarrow 400 \text{ KB/s} \]
what else?

\[ \Rightarrow \text{ small (caches) \small (inside drive) } \]

\[ \Rightarrow \text{ zoned mag. recording} \]

outer tracks: more sectors than inner tracks
Lecture 13

Administrative:

- HW 6: homework problems (virtual memory)

- (optional) final

- midterm
  - hw1 ... 6

- Friday:
  - grade so far

- [ ]