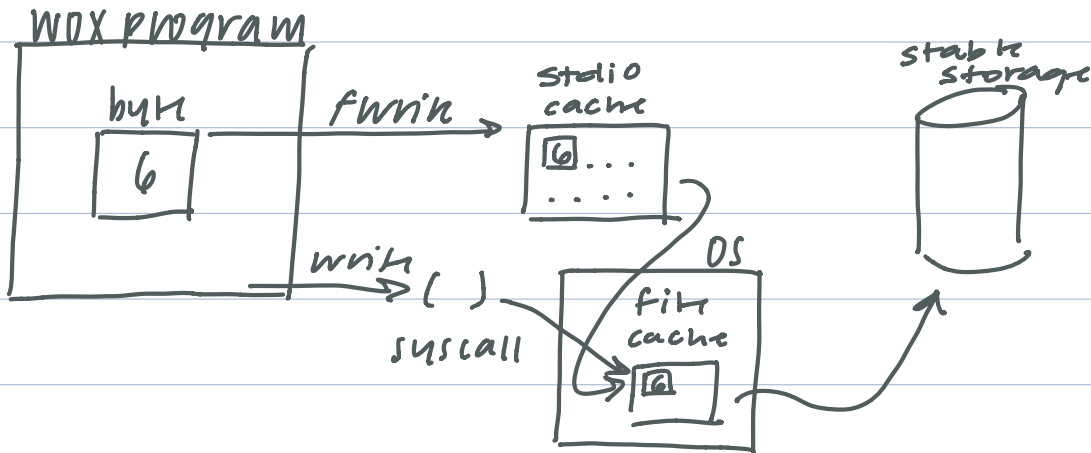


Last Time (writing single byte)



Writing block of text: (512 byte write)
267 times faster

strace

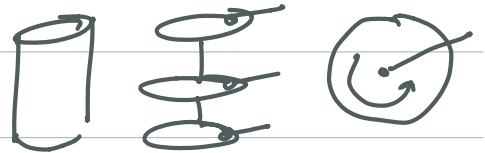
"snoop" on system

Storage Devices

2 aspects of the performance of storage devices

1. latency = time to complete operation (lower is better)

distance, technology effect



2. Throughput: # operations / second (bigger is better)

disk: 5 ms average latency \sim 1 MB/sec

↑ throughput

assumes repositioning head every time

transfer rate \sim 140 MB/second

Disk is high latency, high throughput

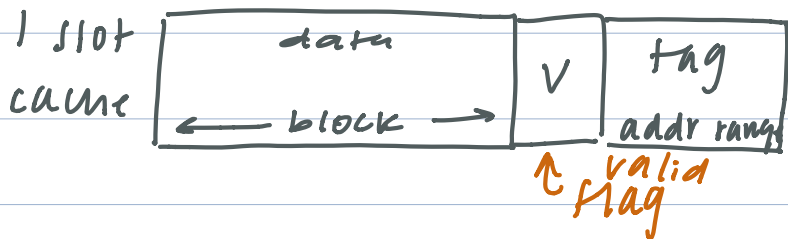
"IO/Hz": may be faster to just get data on hard drive and drive it down IO

	<u>main memory</u>	<u>hard disk</u>	price / mb
2018	\$0.0059 (29.5x)	\$0.000020 (1)	
2003	\$0.09 (4100x)	\$0.00132 (6.6x)	
1990	\$148.20 (7.4Mx)	\$5.45 (270Kx)	
~1995	\$411M (20,500,000,000 000x)	\$6,250	

↑ Moore's law
and silicon,
manufacturing

How does a cache work?

Single Slot



```

request address
↓
if (V) {
  if (reqaddr in $addr range) {
    yes → hit!
    return data
  }
  else {
    no → miss
  }
}

```

Multi-slot cache

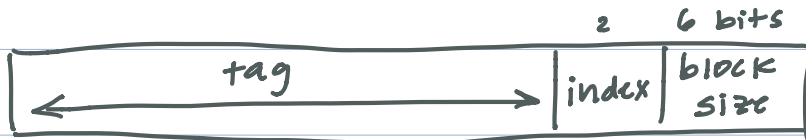
0	data	v	tag
1	data	v	tag
2	data	v	tag
3	data	v	tag

$$C[i] = A[i] + B[i]$$

request addr



fully-associative cache: any member of any addr into any slot of cache



direct mapped: each request addr can only appear in one slot

- in a process, virtual addr
- in hardware, physical addr