ECE 571 – Advanced Microprocessor-Based Design Lecture 31

Vince Weaver

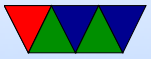
http://web.eece.maine.edu/~vweaver

vincent.weaver@maine.edu

18 November 2020

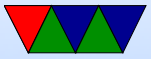
Announcements

- Don't forget HW#10 readings
- Let me know if you want to borrow power measurement devices



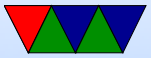
Formatting and 4k Transition

- Low-level. Logical blocks. Delimited by markers start, end, ECC, space to allow for timing
- Traditionally 512B but by moving to 4k can reduce percentage dedicated to the delimiters (though have to be backwards compatible)
- High-level format is just putting filesystem onto the blocks
- Interesting backwards compatibility issues and Linux drama when this first came out



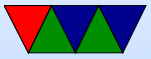
Form Factor

- Capacity – 1GB (10**9) vs 1GiB (2**30)
- Desktop – 60GB - 4TB, 5400 - 10k rpm
- Mobile – smaller in size and space, usually spin slower 5400 or 7200
- Enterprise – fast, often 10k or 15k. Sometimes smaller platters (2.5) for faster seek time
- Consumer – often slower, more shock resistant

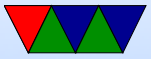


Performance

- Response time vs Throughput
- Seek time – time to get head to block of interest
  Fast drives today, 4ms
- Rotational latency – head seeks to right track but has to wait for block to spin by
  15k – 2ms, 7200 – 4ms, 5400 – 5ms
- Bits per second – 2010, 7200rpm drive 1Gb/s, depends on which track, rpm,.
  SATA can send about 3Gb/s with 10-bit encoding

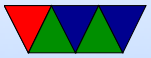


SSD



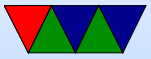
FLASH

- ROM, EPROM, EEPROM
- FLASH (NAND/NOR) can have parts erased, not whole thing at once
- Invented at Toshiba in the 1980s
- NAND vs NOR

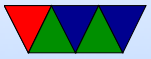


NOR Flash

- Long erase
- Random access
- 100x - 1000000x erase cycles
- CF was originally NOR (but NAND cheaper)
- Like a NOR gate, one end to ground
- low read latency, can be used bit-by-bit ROM
- Program by writing at high voltage, Channel turned on, quantum tunneling via “hot electron injection”
- Resetting (actually to all 1 state). Same process,

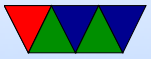


opposite direction, large voltage. Can in theory be individually reset bits but in practice in blocks

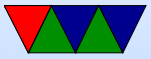


NAND Flash

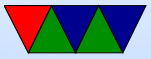
- Reduced erase/write times
- Less chip area (higher density)
- 10x endurance of NOR
- Must read out in large blocks (not random access)
- Wired in series like a bunch of NAND gates
- Certain amount of errors allowed (unlike NOR)
- Tunnel injection for erasing
- ECC error correction
- Programming



  - Starts as all 1s
  - In general can change any 1 to 0 at any time, but if want to switch from 0 back to 1 have to erase whole block
- Floating-gate transistors
  - Each cell like a MOSFET, but with two gates
  - Floating gate and control gate. Control to switch, float can trap electrons
  - Floating gate raises the Vt by acting as a screen, so detect if 1 or 0 by putting an intermediate voltage
- Charge pumps

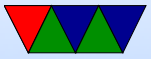


- Need high voltage to write. But usually this is done from single voltage supply
- In space applications usually the charge pump that fails (so chips can still be read, but no longer write)



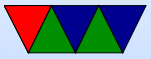
Flash Issues

- Memory wear – can only write so many times before wears out. 100k?
- Memory disturb – a bit like rowhammer, write too many times can change nearby
- Xray can reset bits (problem when trying to see if BGA solder went well)



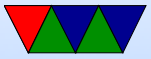
SSD

- Solid-state disk
- No moving parts
- Faster, lower-latency, more resistant to shock
- Still more expensive
- Most 3D TLC NAND-based
  SLC=single level (bit) per core, MLC=multiple(2), TLC=triple level
- SSDs not permanent, will gradually leak and lose data after 2-3 years (faster if worn? trapped electrons leak



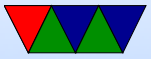
away)

- Originally was DRAM (battery backed) but these days NAND flash
- Controller that handles things
  - Bad-block remapping
  - Read scrubbing
  - Wear leveling



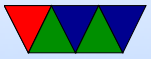
SSD Performance

- DRAM-based fastest
- Single NAND relatively slow
- Having lots in parallel helps

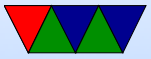


SSD Form Factor

- Can be SATA, but SATA while fast enough for magnetic disk cannot keep up with flash
- M.2 (formerly NGFF) – Intel
  Can provide PCIe, SATA 3.0, or USB 3, different keying to keep from plugging in wrong
- NVMe – non volatile memory express – hook up via PCIe
- U.2 (formerly) SSD Form Factor Working Group, provides SSD connector for enterprise. Hot-swap. mechanically indentical to SATA-express. 3.3V or 12V

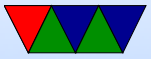


(M.2 only 3.3V)

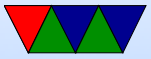


Trim Operation

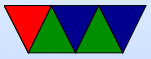
- On filesystem, erase file, usually just mark as deleted and blocks unused, even though never want the bytes again
- TRIM on flash lets you tell disk you don't want them anymore, and the drive can then reclaim them
- Also, when OS then re-uses freed block, flash sees this as an over-write of the block (expensive) rather than a fresh write to a new block
- Expensive because typically erases in big chunks (512kB)



so over-writing you have to erase a whole big chunk, then do a write back of existing values

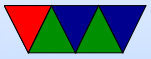


Hybrid Drive

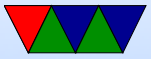


SSD vs HDD comparison

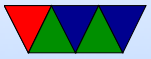
- Data Durability – SSD loses in a few years – HDD lasts longer, but motors/mechanical might fail
- Startup – HDD has to spin up, takes a while
- Random Access – HDD bad, has to spin to location
- Read latency – SSD better
- Bandwidth – SSD often higher
- Read perf – SSD fast but goes down with use
- Noise – SSDs silent
- Heating – both don't like high temps



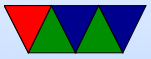
- Cooling – SSDs can operate at lower temps
- Air – SSDs don't require air
- Price – SSD cheaper
- Power –SSD usually better
- Storage size – HDD usually better



Compact Flash

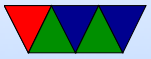


SD-card

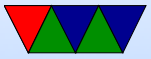


RAID arrays

- Redundant Array of Independent/Inexpensive Disks
- RAID0 – Striping – just makes drive bigger (spreads across disks which might be faster)
- RAID1 – Mirroring, two copies
- RAID2 – bit-level striping not really used anymore
- RAID3 – byte-level striping
- RAID4 – block-level striping
- RAID5 – distributed parity, at least 3 disks, can survive on 2/3 disks



- RAID6 – double parity, any two disks can fail



Tape