Lecture 16: RAID, Files

- RAID

- Filesystem interface
  - files, directories, metadata, links, ...
  - File API
  - VFS
Disks are slow! Getting slower.

- Address by using multiple disks simultaneously.
  - Redundant array of inexpensive disks (RAID)
  - Build big expensive disks (SLED)
    - Single large expensive disk.
Disk read/write in units of sectors (1/2 k - 4 k).
Break up data into blocks to put in sectors
(Analogous to frames/pages).

File 1:
B₁, B₂, B₃, ..., B₁₀

File 2:
B₁, ..., B₄₅

To read File 1 in its entirety, don't use 2nd disk with.

Striping (RAID 0)

Implements performance effectively do sequential reads & writes 2x as fast.
Disks fail:
- Local failures (bad sector)
- Disk failures (entire disk)

With striping: failure is worse.
(As likely for 1 or 2, 3 or 4, 5 or 6)
- either failing => file unreadable

Mirroring (RAID1)
- pro: Great fault tolerance
- con: Compared to striping, not quite as fast. (have to write 2x)

Write speed:
- slower than 1 disk

Read speed:
- almost as fast as striping
RAID 2: bit-level parity

RAID 3: byte-level parity

RAID 4: block-level parity

Parity is the sum of the bits of the other disks, mod 2.

Useful for any combination of the disks.

Parity of $B_{i-3} \oplus 3$, i.e., XOR of 3, 1, 0, 0.

Parity is 1, 1, 1, 1.

$C = B \oplus D \oplus E \oplus F$.

$P = B \oplus D \oplus E \oplus F$

$P = B \oplus D \oplus E \oplus F$

Pro: Fault tolerance: Can lose 1 of $n$ disks, still recover

- Reads as fast as storing across 1 disk

- Parity disk doesn't improve performance

every write to writes to parity disk and old parity, recompute new parity, write new data + new parity.

- To write: Read old data, and old parity, recompute new parity, write new data + new parity.

More wear and (maybe) more seeking.
RAID 5: put parity on different disk for each "row" of blocks.

Pros/Cons:
- Same as RAID 4, except for non-uniform wear.
- Con: requires complicated controller.
Raid 0+1 or 0+1

- Same
- RAID 0
- RAID 0
- RAID 1
- RAID 1
- RAID 0

Pros:
- Good fault tolerance
- Simple
- Good read performance
- Decent write performance

Effectively the same.
RAID 6: Tolerating 1 failure might not be enough.

- Recovery takes a long time for large disks
- Lots of disks, long time = higher prob of 2 simultaneous failures.

- 2 striped parity blocks

- Reed-Solomon codes: generalize parity to handle 2 or more failures (by holding more "parity"-like bits)