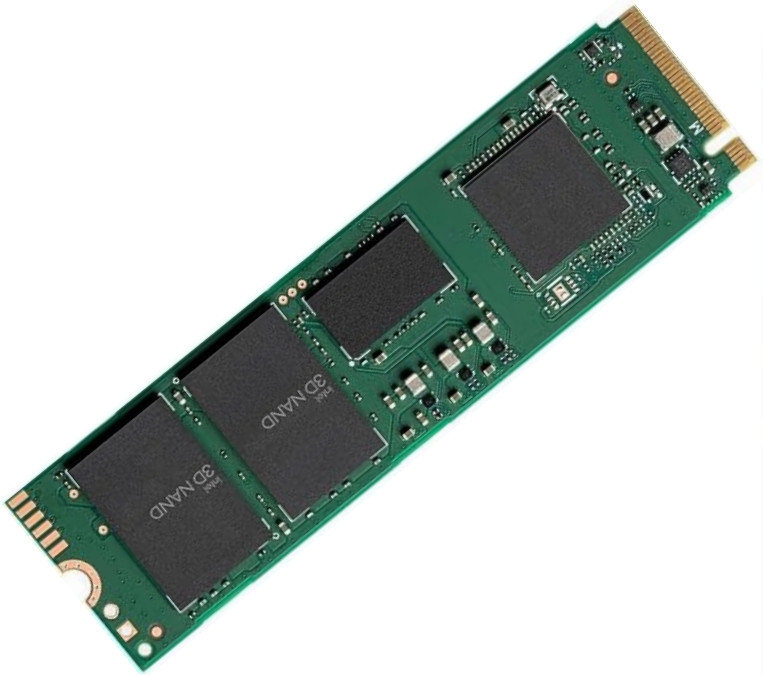
Solid State Drives (SSDs)

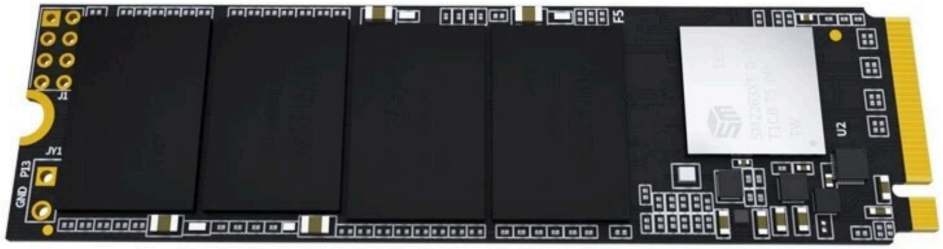
ICS332
Operating Systems

Henri Casanova (henric@hawaii.edu)

Solid State Drives (SSDs)

- Flash-based storage, no moving parts





Advantages over HDDs

- Faster, silent, lower power, more reliable, lighter
- Their market share has been increasing
- The only clear advantage of HDDs **for now**: $ / byte
  - "Greenness"? HDDs are better for the environment to manufacture but much worse to operate…..
- People used to think that SDDs would be the same $/byte cost as HDDs around now, but that hasn't happened yet (it's about 6x)
  - Due in part to new HDD advances
- Some say it will happen by 2030?

SDD Storage Structure

- SSDs store bits into cells
  - Each cell can store 1, 2, or 3 bits depending on the technology
- Cells are organized into **pages** (e.g., 4KB)
- Pages are grouped into **blocks** (e.g., 128KB, 256KB)
- Blocks are grouped into banks (or planes)

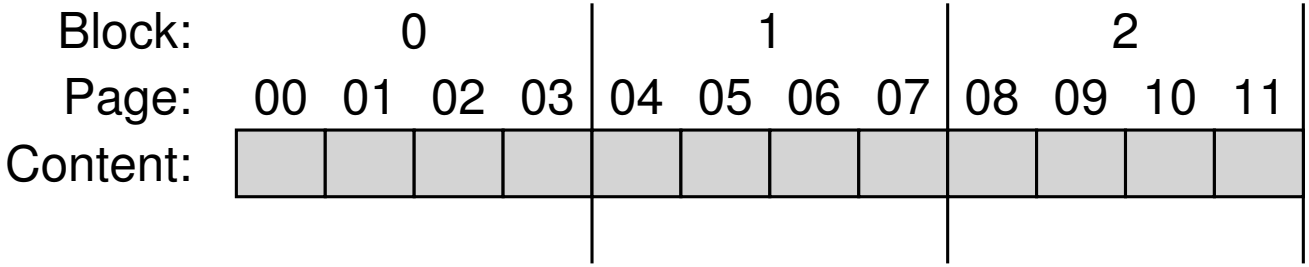


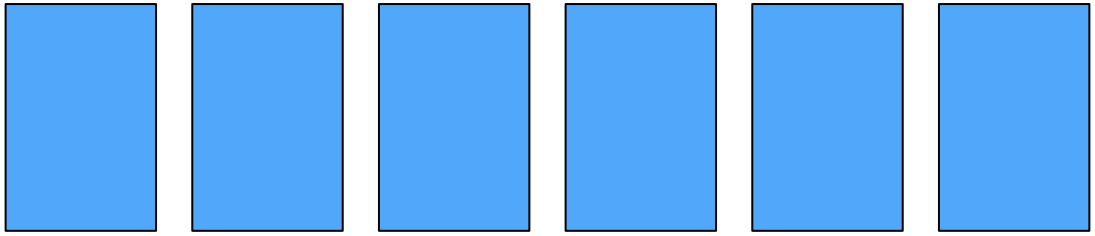
Figure 44.1: **A Simple Flash Chip: Pages Within Blocks**

SSD Operations

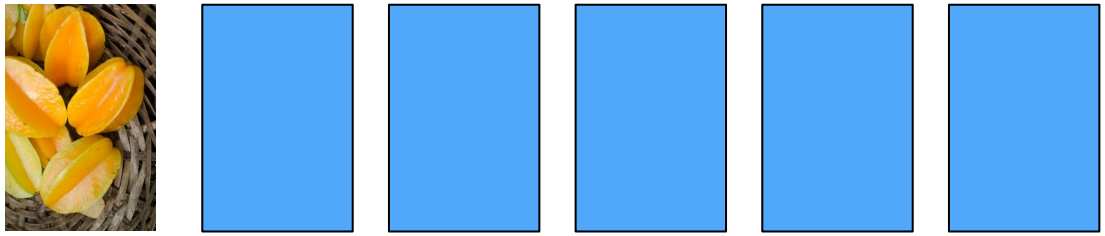
- **Read a page**
  - Very fast (µs), random access makes no difference (the major advantage over HDDs: locality doesn't matter)
- **Erase a block**
  - Much more expensive (ms)
- **Write (a.k.a. *program*) a page**
  - Requires that the page's block has been erased first!!!!
    - This is the "SSD weirdness": To update data in a page, you need to erase the whole block of pages
  - To make things worse: this causes wear out of the flash cells
  - Other problem: if you want to update, say, only one of the pages in a block, you need to first copy all other pages somewhere (e.g., the SSD's controller, RAM), then erase the whole block, and then write to all pages
  - This is called write amplification (we wrote more data than needed)
- Let's see this on an example….

Write Amplification Example (1)

- Say we have a 6-page block (each page is 4KB)

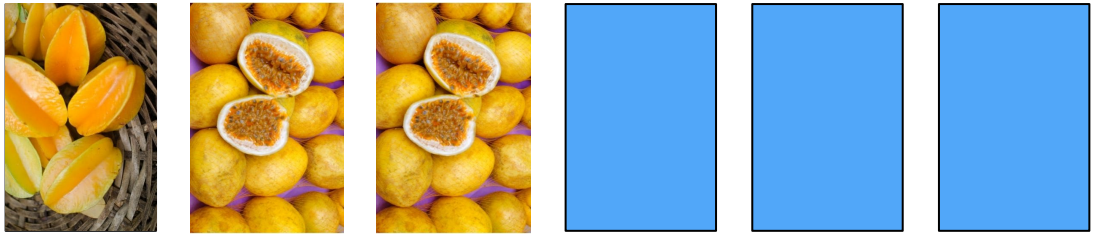


- Let's write a 4KB file



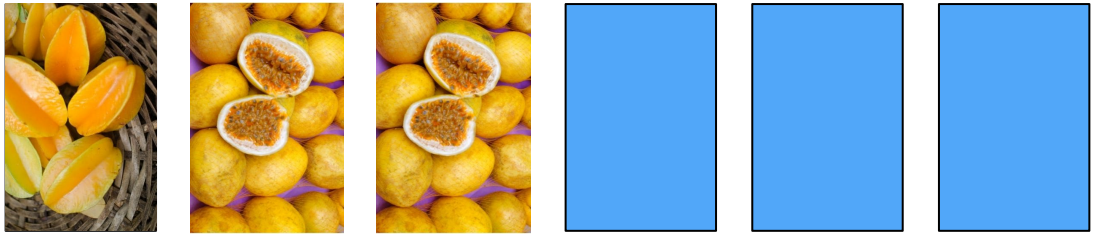
Write Amplification Example (2)

- Let's write a 8KB file

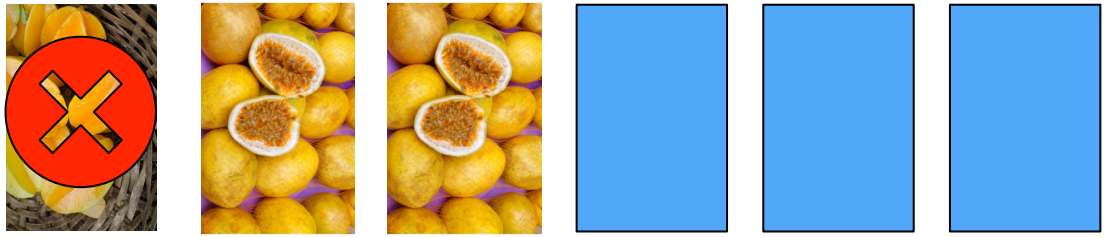


Write Amplification Example (2)

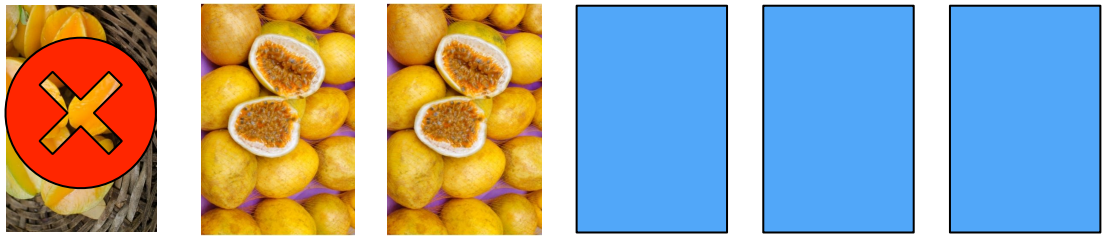
- Let's write a 8KB file



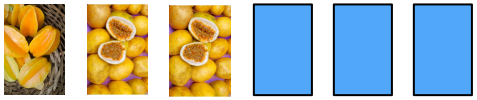
- Let's erase the first file
  - Instead of erasing the whole block, the SSD controller just marks the first page as invalid



Write Amplification Example (3)



- Now we want to write a 16KB file 
- The whole block's data is loaded into the SSD's controller's cache



- The data is updated in the cache

- The data is written to the block on the device

Write Amplification Example (4)

- We wanted to write 4KiB + 8KiB + 16KiB = 28KiB of application data
- We had to write 4KiB + 8KiB + 24KiB = 36KiB of data to the SSD
- We could come up with an example where we write about 5x more than what we need to write
- For this reason, the controller keeps writing on the SSD until full, before it attempts any rewrite
- Once full, writes are then more and more amplified
  - i.e., there is fragmentation everywhere, and no “free” blocks where data can be written
- And rewriting a block over and over leads to a wear out
- In the end, performance is still very good relative to that of an HDD, in part because SSDs employ several techniques….

SSD Techniques

- **Increase Performance:** the controller can clean up blocks with invalid pages at any time so that they're easily writable later
  - This is called garbage collection!
  - Same idea as the OS writing back dirty pages to disk once in a while when the disk is idle so that at the next page fault there is no need to write the page back
- **Decrease wear out:** the controller tries to spread writes over all pages, so that they would all wear out together as late as possible
  - In practice, for most "normal" workloads, wear out isn an issue
  - i.e., your SSD will be ok until you get your next laptop unless you do something very unusual

Main Takeaways

- SSDs are better than HDDs in almost all respect besides cost per byte
- The main issue: write amplification
- Techniques have been developed to deal with it (and with wear-and-tear), and nowadays these problems are no longer a big issue in practice

Conclusion

- SSDs are here to stay
  - By now we all have one!
- They are still more expensive than HDDs
  - HDDs will continue being used to store large dataset for the upcoming years
  - Just like tape storage is still in use!
- OSTEP has many more details in Chapter 44