Chapter 44: Flash-based SSDs

1. Introduction to SSDs

- **Solid State Drives (SSDs)** are storage devices that use flash memory to store data.
- They have no moving parts, which makes them much faster than traditional hard disk drives (HDDs).
- The core technology is NAND-based flash memory.

2. Flash Memory Fundamentals

- **Hierarchy:** Flash memory is organized into a hierarchy:
  - **Page:** The smallest unit that can be read or written (e.g., 4KB).
  - **Block:** A collection of pages (e.g., 256KB).
  - **Plane:** A collection of blocks.
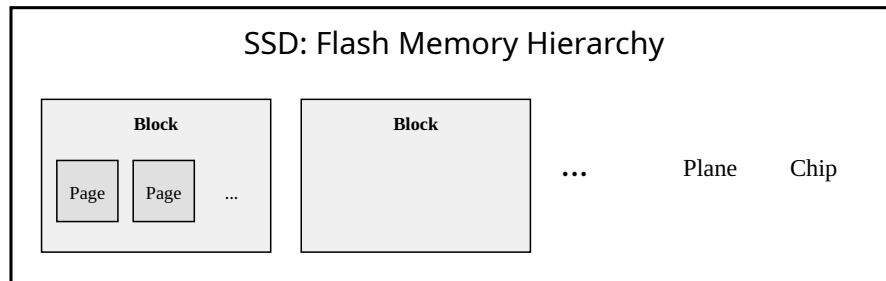 - **Chip:** A collection of planes.

Figure 1: SSD Hierarchy

- **Operations:**
  - **Read:** Reading a page is very fast.
  - **Program (Write):** Writing to a page is slower than reading.
  - **Erase:** Before a page can be written, the entire block it belongs to must be erased. Erasing is a very slow operation.
- **The Write Problem:** You cannot overwrite a page directly. You must erase the entire block first. This makes writing to flash memory more complex than writing to a hard drive.

3. The Flash Translation Layer (FTL)

- The **Flash Translation Layer (FTL)** is a layer of software that runs on the SSD controller.
- It hides the complexities of flash memory from the operating system and presents a standard block device interface.
- **Key Responsibilities:**

- **Logical to Physical Mapping:** Maps logical block addresses (LBAs) from the OS to physical page addresses (PPAs) on the flash chip.
- **Garbage Collection:** Reclaims blocks that contain invalid data.
- **Wear Leveling:** Distributes writes evenly across all blocks to prevent them from wearing out prematurely.
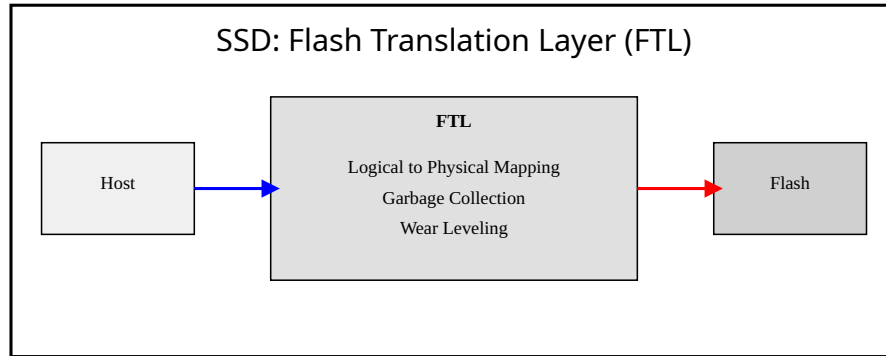

Figure 2: FTL

4. Log-structured FTL

- Most modern FTLs are **log-structured**.
- **Key Idea:** Instead of overwriting data in place, new data is always written to the next available free page, like a log.
- This avoids the costly erase-before-write cycle for every write.
- The FTL maintains a mapping table to keep track of the physical location of each logical block.

5. Garbage Collection

- When data is updated, the old version becomes invalid ("garbage").
- The FTL's **garbage collector** reclaims blocks containing garbage by copying any live data to a new block and then erasing the old block.
- The `TRIM` command allows the OS to inform the SSD which blocks are no longer in use, which helps the garbage collector to be more efficient.

6. Write Amplification

- **Write amplification** is the ratio of data written to the flash chips to the data written by the host computer.
- A high write amplification can hurt performance and reduce the lifespan of the SSD.
- It is caused by factors such as garbage collection and small writes.
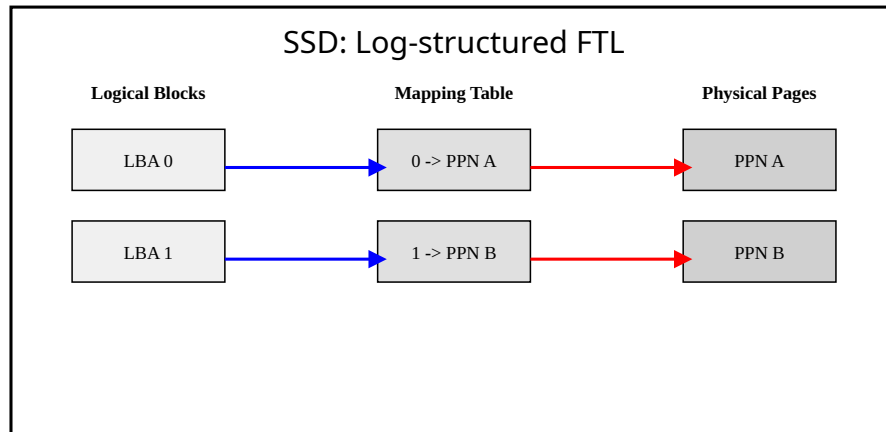
Figure 3: Log-structured FTL

7. SSD vs. HDD

- **Performance:** SSDs are much faster than HDDs, especially for random I/O.
- **Cost:** SSDs are more expensive than HDDs on a per-gigabyte basis.
- **Hybrid Systems:** Many systems use a hybrid approach, with a small, fast SSD for the OS and frequently accessed data, and a large, cheap HDD for bulk storage.