NAND Flash Memory Technology: The Basics of a Flash Memory Cell

by Zachary Painter April 12, 2018 0 Comment

If you google NAND Flash Memory Technology, you’re probably looking for very specific information explaining how NAND Flash works, or what sets it apart from other types of Flash memory.

But it only takes one cursory look over the first page of Google’s search engine results to bewilder you: Terms like “bits per cell” or “floating gate” appear and you start to feel out of your depth.

The truth is, learning about NAND Flash is easier if you know where to start your research. But we’ve taken care of that step for you. Read on to learn more about the basics of NAND Flash and the technology that sustains your SSD.

Table of Contents:

1. What Exactly Is NAND Flash Memory?
2. NAND Chip Architecture
3. NOR Flash vs. NAND Flash
4. Floating Gate Transistors
5. Storing Your Data

What Exactly Is NAND Flash Memory?

Flash memory is a non-volatile solid-state storage medium that relies on electric circuits to store and retrieve your data. What makes Flash memory “non-volatile” — or capable of maintaining your data without power — boils down to the underlying technology at work in a Flash Memory Cell.

To better understand what Flash memory cells are, we should first discuss NAND Flash chips.

NAND Flash Chip Architecture

NAND Flash chips are the black, rectangular blocks attached to the Printed Circuit Board (PCB) of your SSD (or any Flash-based product). Not to be confused with the DRAM
Cache or Controller, NAND Flash chips are usually consolidated to one area of the PCB as shown below.

NAND Flash memory cells are housed within NAND chips. The layout of a NAND chip — or “die” — consists of four areas:
Die: Semiconductor wafers are cut from a block of electronic-grade silicon. The piece of wafer cut from the block is referred to as a die.

Plane: One die contains one to two planes. Planes can (generally) execute identical operations together.

Block: Planes contain varying numbers of blocks. NAND Flash cells can only be erased at the block level.

Page: Blocks usually contain around 128 pages. Pages are the smallest units that can be programmed (or written to).

**How many cells constitute a page?**

Each page contains anywhere from **512 – 4,096 bytes** (1 byte = 8 bits). Some cells hold just one bit, while others contain up to three depending on the SSD.

Now that we know how NAND Flash cells are organized, how does NAND work?

**NOR Flash vs. NAND Flash**

There are two main types of Flash memory: **NOR** and **NAND**. Both types integrate a grid system of Flash memory cells that are wired together. How these cells are wired together differentiates the two.
NOR and NAND wiring

Cells are organized in an array of rows and columns (think of the grid system we just mentioned). In both NOR and NAND configurations, word lines are the wires connecting the rows, and bit lines are the wires connecting the columns.

What are word lines and bit lines?

Earlier, we said Flash memory relies on electric circuits to store and retrieve data. Well, the intersecting word lines and bit lines are a part of that network of circuits. Word lines assess the cell’s electrical charge, and bit lines regulate the flow of electrical current from cell to cell.

Before we go any further, let’s view the schematics below to get an idea of what NOR and NAND look like:

In the diagram above, we see the bit lines run alongside the cells (connecting the columns) and the word lines connect the rows. The key detail here is that each cell requires a word line connector and bit line connector.
NAND Flash, on the other hand, links an entire cell column by running the bit line *through* each cell, bypassing the need for an extra bit line connector on the side.

**So, what does the wiring really mean?**

What’s important here is that, due to their wiring configurations, NOR offers better read speeds and random access capabilities; NAND offers better write/erase capabilities with a cheaper cost per bit, so it’s better suited for the data demands of SSDs, Flash Drives, and Flash Memory Cards.

Curious about SSDs? Learn more about what SSDs are and why you need one for your PC or laptop!

**Floating Gate Transistors**

Each NAND Flash memory cell contains a Floating Gate Transistor, which is where the Program/Erase cycle of a NAND Flash cell takes place.
The floating gate transistor consists of **five main elements:**

- Control Gate
- Floating Gate
- Insulator, or “Oxide Layer”
- P-Substrate (Poly-Silicon)
- Source and Drain

In a standard NAND configuration, word lines plug into the transistor’s **control gate** (horizontally across rows) and bit lines link each cell’s **source** and **drain** (vertically along columns). The source is where electrical current enters the cell, and the drain is where it exits.

**Here’s how it works:**

The control gate opens and closes from voltage applied by the word lines. More specifically, the word line checks to see if the cell holds a charge (a “0”) or not (a “1”), also known as the cell’s **bit state**.

A cell registering as a “1” (no charge) indicates there are **no electrons present** in the **Floating Gate (FG)**; if the cell’s bit state reads as a “0” (holding a charge), there **are** electrons in the FG.

**How do electrons get into and out of the FG?**

By applying a positive charge to the bit line and word line, electrons stored in the **source** get excited and plow through the **P-Substrate** field to the **drain**.
When this happens, some electrons squeeze through the “insulated” layer into the floating gate, a process called **tunnel injection**. Forcing electrons into a floating gate is how your drive writes data to a cell.

Writing — or **programming** — happens at the **byte level**, and sets individual bits from “1” to “0”.

Erasing data happens at the **block level**. A similar process called **tunnel release** forces electrons out of the floating gate, in turn changing the bit state to “1”.

Electrons stored in the floating gate are unaffected by the absence of power, which is how Flash technology achieves and maintains its state of non-volatility.

**Storing Your Data**

Your data breaks down into a series of ones and zeroes stored in Flash memory cells. These binary digits, as we’ve already discussed, are called bits.

An SSD’s capacity is determined by how many bits it can store per cell, hence the terms **Single Level Cell (SLC)**, **Multi Level Cell (MLC)**, and **Triple Level Cell (TLC)**.

Of the three types of SSDs, SLC is the top performer, followed by MLC and TLC. It’s the fastest drive and the most reliable.
Why?

SLC has the least cell density, so operations like programming and erasing are less complex. For example, when a cell contains one bit, its bit state can only read as a 1 or 0.

By introducing multiple bits per cell, a cell’s bit state could have as many as eight possibilities, creating a larger margin for error.

The advantages of higher cell density are increased capacity at a cheaper cost per bit, so MLC and TLC are the preferred consumer level drives.

SLC is a low maintenance drive, but this has its own trade-offs. Check out our SLC, MLC, and TLC guide to learn more about NAND Flash and its effects on drive performance.

Is there anything else to know?

There is always more ground to cover when discussing NAND Flash. Stay tuned for our future guides covering standard SSD maintenance technologies (like Bad Block Management, Wear Leveling, and Error Correcting Code) to learn more!

If you have any questions about NAND Flash technology, or SSDs in general, feel free to leave a comment below!