USB FLASH DRIVE

To start off we are going to have a very basic course in electricity. This is necessary because we all come from such a diverse background that we need a common foundation. From that I can then build a discussion on flash drives, and we will all be together. So you electronic engineers bear with the rest of us while we catch up with you in 90 seconds.

Electricity is simple to understand if we consider it is nothing but the moving of electrons. Electrons are actually nothing more than negative charges. These negative charged electrons are attracted to a positive charge and will do almost anything to get to the positive charge. In the use of electricity we put stuff in the path of these electrons, causing them to work for us. Run motors, light light bulbs or run computers. Normally the electrons travel on wires, but in are quest to make them do some things we sometimes put them in various other mediums. One thing we can always count on is that they will take the path of least resistance.

We can put all sorts of objects in the way of this path to regulate the amount of electron flow. This flow is called current. The restriction to the flow is called resistance and takes the form of resistors. When the resistance become so great that it allows no electron flow it is called insulation. However, no insulation is so great that electrons if pushed hard enough will jump this insulation. Example sparks, lightning.

You now know in 90 seconds as much as the electrical engineers, that spent at least 4 years to learn. Let’s take this very rudimentary knowledge of electronics and put it to use in making flash memory. The structure we use in flash memory is called floating gates. A floating gate is a type of a transistor. Transistors have been around most of our adult life. Prior to the transistor we used vacuum tubes. A transistor is like a faucet controlling water. If you can think of the faucet regulating the amount of water flow by electronically changing the faucet knob you have all you need to know about transistors. Floating gates are a little different than a regular transistor or faucet. In a floating gate we force electrons across an insulator and leave them deposited on a fully insulated plate. Thus the name floating gate. The width of the Floating Gate transistor will average between 12 and 25 nanometers, depending on the process technology used and the density of the number of flash cells on a chip. A nanometer is only three to five atoms wide. They're small. Really small. ~40,000 times smaller than the width of an average human hair. The features are built through multiple uses of photolithography, thin film deposits, and ion implantation. We will go into that later.

Place an electronic charge along the flow of electrons so that the flow of electrons is blocked or allowed to flow depending on the charge. The barrier is usually some type of oxide which surrounds the floating gate entirely, therefore trapping a charge on the floating gate.

On top is the control gate which is used to place the charge on the floating gate. Below this is the floating gate insulated all around by an oxide layer. Below the floating gate is the channel. Because the Floating Gate electrically isolated by its insulating layer, any electrons placed on it are trapped there and, under normal conditions, will not discharge for many years. During read-out, the channel will become conducting or remain insulating, depending on the charge of the cell, which is in turn controlled by charge on the Floating Gate. The current flow through the channel is sensed and forms a binary code, reproducing the stored data.

Let me stop and let everyone get together again. What is binary code? I like to think of it as Morse code

Remember in Morse code we represent a letter by dots and dashes. We don't have the luxury of dashes in computer code so we represent letters by turning the electron flow off and on. If we do this at a given frequency we can write what is known as binary code. This takes a lot more code to write a letter, but we can do it at astronomical frequency.

Back to Floating Gates. A gate will be activated by either placing electrons on the floating gate or not. Thus the gate once programmed remembers only that is off or on. It will then take 8 gates to represent one letter. Therefore if you write a gigabyte of information on a flash drive it will use a billion floating gates. Now don't all of you engineers start correcting me. This is a simplistic example and I am taking some liberties with terms.

There is a bunch of different types and methods that use this floating gate technology. For our purpose I will only go into two methods that these floating gates are packaged. One is called Single-Level Cell Flash designated as SLC, and the other is Die-Stacking or Multi-level. In order to economically increase the amount of bit-storage that a Flash memory chip can accommodate, manufacturers often utilize die-stacking technologies. These technologies result in a Flash memory chip having the capability to store more data in a single chip, but sacrifice speed.

When a Flash storage device is manufactured, steps are taken to ensure that the device operates reliably and to permit the computer to access the memory cells — i.e., to store and retrieve data on the Flash storage device.

These steps — loosely called "formatting" — utilize some of the memory cells within the device and thus reduce the capacity available for data storage by the end user.

Formatting includes the following operations:

1. Testing each memory cell in the Flash storage device.
2. Identifying all defective cells and taking steps to ensure that no data will be written to or read from a defective cell.
3. Reserving some cells to serve as "spares." Flash memory cells have a long but finite lifetime. Therefore, some cells are held in reserve to replace any memory cells that may fail over time.
4. Creating a File Allocation Table (FAT) or other directory.

- Flash storage devices have no moving parts and thereby are not subject to the mechanical failure issues of hard drives. Their overall data reliability enabled them to dominate the portable memory market, operating silently with a zero decibel noise level.

All flash drive memory is NOT CREATED EQUAL!

Let's look at how flash drive memory chips are made, graded and sold. There are 4 different classes and standards.

Grade A:
Brand Name. These are bona fide premium chips with the flash chip manufacturers name and serial number imprinted on the chip itself. They are the most expensive, but also are the most reliable memory chips you can buy. A flash drive supplier using these chips will readily offer the customer a lifetime warranty.

Grade B:
OEM flash chips. These are very similar chips as found in the Brand Name category. The only difference is that the manufacturers put their clients name on the chips. These chips may be just as reliable but do not come with the same warranty as brand name chip.

Grade C:
This is where the problems begin. Often called recycled or reclaimed flash chips, these chips are not recycled chips but are considered to be waste from the silicone wafer that the original manufacturer does not want and considers to be garbage. They are discarded and sold by the pound.
A silicone wafer is a thin small round disk. The center square contains the prime chips. These are the Grade A chips. They perform flawlessly.
The outer edges of the silicone wafer contain chips that have flaws and failures. These are sold to toy manufacturers; cheap gadget makers etc as well as unscrupulous flash drive makers. These waste memory flash drives are found in the least expensive flash drives advertised on-line.

We could stop right here, but I think the most exciting part of Flash Memory is how they make transistors that are only 12 to 25 nanometers wide, less than a thousand of the width of a hair.

Along comes Silicon. More than 90% of the earth's crust is composed of Silica, making silicon the second most abundant element on earth. When sand glitters in sunlight, that's silica. Most importantly to technology, silicon is the principle platform for semiconductor devices. The most advanced semiconductor technologies of today and tomorrow require monocrystalline Silicon with precise uniform chemical characteristics.

What makes silicon so perfect for semiconductor use is it ability to allow electrons to flow in one direction while forming as an insulator in the other direction? Note that in this slide the atoms of the crystal are lined up perfectly in a front to back configuration. The next slide if the atomic arrangement of an Amorphous or non crystal substance which depending on the material allows no electron flow in any direction or electron flow in all directions.

The process to transform raw silicon into a useable single-crystal substrate for modern semiconductor processes begins by mining for relatively pure Silicon Dioxide. Most silicon now is made by reduction of Silicon Dioxide with Carbon in an electric furnace. With carefully selected pure sand, the result is commercial brown silicon of 97% purity or better. This is the silicon eventually used for semiconductors, but it must be further purified to bring impurities below the parts-per-billion level.

When the high level of purity has been attained, the atomic structure of the silicon must be dealt with. A process known as Crystal Growing transforms this polycrystalline silicon into samples with a singular crystal orientation. The Polysilicon is mechanically broken into 1 to 3 inch chunks and undergoes stringent cleaning in a clean room environment. These clean rooms are a thousand time cleaner than the floor of a hospital. Which from what we read may not be all that clean. These chunks are then packed into quartz crucibles for meltdown (at 1420oC) in a furnace. A monocrystalline silicon seed is installed into a seed shaft in the upper chamber of the furnace. Slowly, the seed is lowered into the silicon melt. Next, the seed is slowly retracted from the surface allowing the melt to solidify at the boundary. As the seed pulls the silicon from the melt, both the crucible and the seed are rotated in opposite directions
to allow for an almost round crystal to form. The furnaces also must be very stable and isolated from vibrations. Once the proper crystal diameter is achieved, the seed lift is increased. This, along with the heat transfer from heater elements will control the diameter of the crystal. Once the growth process is complete, the crystal is cooled inside the furnace for up to 7 hours. This gradual cooling allows the crystal lattice to stabilize and makes handling easier before transport to the next operation. Ingots coming from crystal growing are slightly over-sized in diameter and typically not round. Hence, a machine employing a grinding wheel shapes the ingot to the precision needed for wafer diameter control. Other grinding wheels are then used to carve a flat on one side in order to define the proper orientation of the future wafer to a particular crystal axis.

Wafer shaping involves a series of precise mechanical and chemical process steps that are necessary to turn the ingot segment into a functional wafer. It is during these steps that the wafer surfaces and dimensions are perfected to exacting detail. Each step is designed to bring the wafer into compliance with each customer specification. The first of these critical steps Multi-Wiring Slicing. Here, a thin wire is arranged over cylindrical spools so that hundreds of parallel wire segments simultaneously travel through the ingot.

State of the art front surface polishing is performed generally in a two step process. One mechanical polishing step to create flatness. Second, lapping the wafers removes saw marks and surface defects from the front and backside of the wafers, thins the wafer to specifications and relieves much of the stress accumulated in the wafer during the sawing process. This is followed by a chemical etch to create smoothness.

For many applications, the quality of a polished wafer is not sufficient. This is mainly due to defects generated during crystal growth in the bulk of the wafer. These defects, when they are within a few microns to the surface, can deteriorate the performance of devices built on top. Presently, the best solution to this problem is to deposit an additional layer of high purity Silicon on the top of a polished wafer substrate.

Two sizes of wafers are produced in a fab lab. The size that is being passed around is approximately 7 ¾ inches containing 88 die, and a larger size that is approximately 11 ¾ inches and contains 232 die
After a final clean and polish; wafers are ready for a final inspection before delivery.