

Non-volatile memories are used to maintain system state even when the system is powered down. In flash memories, long-term storage is achieved by storing charge on an well-insulated conductor called a floating gate, where it will remain stable for years. The floating gate is incorporated in a standard MOSFET, placed between the MOSFET's gate and the MOSFET's channel. If there is no charge stored on the floating gate, the MOSFET can be turned on, i.e., be made to conduct, by placing a voltage V_1 on the gate terminal, creating an inversion layer that connects the MOSFET's source and drain terminals. If there is a charge stored on the floating gate, a higher voltage V_2 is required to turn on the MOSFET. By setting the gate terminal to a voltage between V_1 and V_2 , we can determine if the floating gate is charged by testing to see if the MOSFET is conducting. In fact, if we can measure the current flowing through the MOSFET, we can determine how much charge is stored on the floating gate, making it possible to store multiple bits of information in one flash cell by varying the amount of charge on its floating gate. Flash cells can be connected in parallel or series to form circuits resembling CMOS NOR or NAND gates, allowing for a variety of access architectures suitable for either random or sequential access. Flash memories are very dense, approaching the areal density of DRAMs, particularly when each cell holds multiple bits of information. Read access times for NOR flash memories are similar to that of DRAMs, several tens of nanoseconds. Read times for NAND flash memories are much longer, on the order of 10 microseconds. Write times for all types of flash memories are quite long since high voltages have to be used to force electrons to cross the insulating barrier surrounding the floating gate. Flash memories can only be written some number of times before the insulating layer is damaged to the point that the floating gate will no longer reliably store charge. Currently the number of guaranteed writes varies between 100,000 and To work around this limitation, flash chips contain clever address mapping algorithms so that writes to the same address actually are mapped to different flash cells on each successive write. The bottom line is that flash memories are a higher-performance but higher-cost replacement for the hard-disk drive, the long-time technology of choice for non-volatile storage.