Semiconductor Memories

Contributed by
Seung-Moon Yoo

10.1 Introduction

Semiconductor memory arrays capable of storing large quantities of digital information are essential to all digital systems. The amount of memory required in a particular system depends on the type of the application, but, in general, the number of transistors for the information (data) storage function is much larger than the number of transistors used for logic operations and other purposes. The ever-increasing demand for larger data storage capacity has driven the fabrication technology and memory development toward more compact design rules and, consequently, toward higher data storage densities. Thus, the maximum realizable data storage capacity of single-chip semiconductor memory arrays approximately doubles every two years. On-chip memory arrays have become widely used subsystems in many VLSI circuits, and commercially available single-chip read/write memory capacity has reached 1 gigabits (1 Gb). This trend toward higher memory density and large storage capacity will continue to push the leading edge of digital system design.

The area efficiency of the memory array, i.e., the number of stored data bits per unit area, is one of the key design criteria that determine the overall storage capacity and, hence, the memory *cost per bit*. Another important issue is the memory access time, i.e., the time required to store and/or retrieve a particular data bit in the memory array. The access time determines the memory *speed*, which is an important performance criterion of the memory array. Finally, the static and dynamic *power consumption* of the memory array is a significant factor to be considered in the design, because of the increasing importance of low-power applications. In the following, we will investigate different types of MOS memory arrays and discuss in detail their operations and design issues related to area, speed, and power consumption for each type.

The semiconductor memory is generally classified according to the type of data storage and data access. Read/write (R/W) memory must permit the modification (writing) of data bits stored in the memory array, as well as their retrieval (reading)

2)

on demand. The read/write memory is commonly called *Random Access Memory* (RAM), mostly due to historical reasons. Unlike sequential-access memories such as magnetic tapes, any cell can be accessed with nearly equal access time. The stored data is volatile; i.e., the stored data is lost when the power supply voltage is turned off. Based on the operation type of individual data storage cells, RAMs are classified into two main categories. *Dynamic RAMs* (DRAM) and *Static RAMs* (SRAM). The DRAM cell consists of a capacitor to store binary information, 1 (high voltage) or 0 (low voltage), and a transistor to access the capacitor. Cell information (voltage) is degraded mostly due to a junction leakage current at the storage node. Therefore, the cell data must be read and rewritten periodically (*refresh operation*) even when memory arrays are not accessed. On the other hand, the SRAM cell consists of a latch, therefore, the cell data is kept as long as the power is turned on and refresh operation is not required. Due to the advantage of low cost and high density, DRAM is widely used for the main memory in personal and mainframe computers, and engineering workstations. SRAM is mainly used for the cache memory in microprocessors, mainframe computers, engineering workstations, and memory in hand-held devices due to high speed and low power consumption.)

(*Read-Only Memory* (ROM) allows, as the name implies, only retrieval of previously stored data and does not permit modifications of the stored information contents during normal operation. ROMs are nonvolatile memories; i.e., the stored data is not lost even when the power supply is off and refresh operation is not required. Depending on the type of data storage (data write) method, ROMs are categorized as *Mask ROM,* in which data are written during chip fabrication by using a photo mask, and *Programmable ROM* (PROM), in which data are written electrically after the chip is fabricated. Depending on data erasing characteristics, PROMs are further classified into *Fuse ROM, Erasable PROM* (EPROM) and *Electrically Erasable PROM* (EEPROM). The data written by blowing the fuse electrically cannot be erased and modified in Fuse ROM. Data in EPROMs and EEPROMs can be rewritten, but the number of subsequent re-write operations is limited to 10^4–10^5. In EPROMs, ultraviolet rays that can penetrate through the crystal glass on the package are used to erase whole data in the chip simultaneously, while high electrical voltage is used to erase data in 8 bit units in EEPROMs. *Flash memory* is similar to EEPROM, where data in the block can be erased by using a high electrical voltage. A drawback of EEPROM is the slower write speed, in the order of microseconds.)

Ferroelectric RAM (FRAM) utilizes the hysteresis characteristics of a ferroelectric capacitor to overcome the slow write operation of other EEPROMs. ROMs are generally used for permanent (look-up) memory in printers, fax, and game machines, and ID cards due to lower cost than RAM. Figure 10.1 and Table 10.1 show an overview and characteristic summary of the different memory types and their classifications, respectively.

Figure 10.2 shows equivalent circuits of memory cells. The DRAM cell consists of a capacitor and a switch transistor. The data are stored in the capacitors as presence and absence of charge: The presence of charge in the capacitor is considered as data "1" while the absence of charge in the capacitor as data "0". The stored charge is subject to gradual decay due to leakage current, thus *refresh* operation is required.