Once programmed, the charge on the floating gate cannot be removed electrically. UV photons cause the dielectric to become slightly conductive, allowing the floating gate's charge to gradually drain away to its unprogrammed state. This UV erasure feature is the reason why many EPROMs are manufactured in ceramic packages with transparent quartz windows directly above the silicon die. These ceramic packages are generally either DIPs or PLCCs and are relatively expensive. In the late 1980s it became common for EPROMs to be manufactured in cheaper plastic packages without transparent windows. These EPROM devices are rendered one-time programmable, or OTP, because it is impossible to expose the die to UV light. OTP devices are attractive, because they are the least expensive nonmask ROM technology and provide a manufacturer with the flexibility to change software on the assembly line by using a new data image to program EPROMs.

The industry standard EPROM family is the 27xxx, where the "xxx" indicates the chip's memory capacity in kilobits. The 27256 and 27512 are very common and easily located devices. Older parts include the 2708, 2716, 2732, 2764, and 27128. There are also newer, higher-density EPROMs such as the 27010, 27020, and 27040 with 1 Mb, 2 Mb, and 4 Mb densities, respectively. 27xxx EPROM devices are most commonly eight bits wide (a 27256 is a 32,768 × 8 EPROM). Wider data words, such as 16 or 32 bits, are available but less common.

Older members of the 27xxx family, such as early NMOS 2716 and 2732 devices, required 21-V programming voltages, consumed more power, and featured access times of between 200 and 450 ns. Newer CMOS devices are designated 27Cxxx, require a 12-V programming voltage, consume less power, and have access times as fast as 45 ns, depending on the manufacturer and device density.

EPROMs are very easy to use because of their classic asynchronous interface. In most applications, the EPROM is treated like a ROM, so writes to the device are not an issue. Two programming control pins, V_{PP} and PGM*, serve as the high-voltage source and program enable, respectively. These two pins can be set to inactive levels and forgotten. What remains are a chip enable, CE*, an output enable, OE*, an address bus, and a data output bus as shown in Fig. 4.3, using a 27C64 (8K × 8) as an example.

When CE* is inactive, or high, the device is in a powered-down mode in which it consumes the least current—measured in microamps due to the quiescent nature of CMOS logic. When CE* and OE* are active simultaneously, D[7:0] follows A[12:0] subject to the device's access time, or propagation delay. This read timing is shown in Fig. 4.4.

When OE* is inactive, the data bus is held in a high-impedance state. A certain time after OE* goes active, t_{OE}, the data word corresponding to the given address is driven—assuming that A1 has been stable for at least t_{ACC}. If not, t_{ACC} will determine how soon D1 is available rather than t_{OE}. While OE* is active, the data bus transitions t_{ACC} ns after the address bus. As soon as OE* is removed, the data bus returns to a high-impedance state after t_{OEZ}.

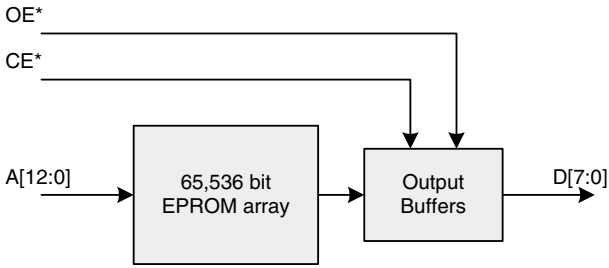

FIGURE 4.3 27C64 block diagram.

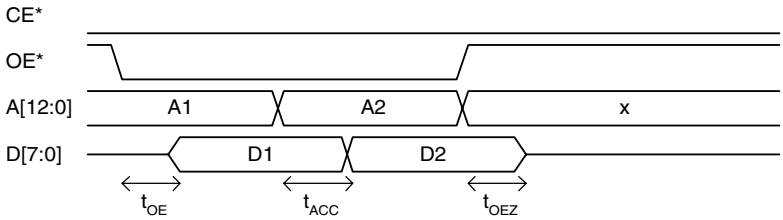

FIGURE 4.4 EPROM asynchronous read timing.

Many microprocessors are able to directly interface to an EPROM via this asynchronous bus because of its ubiquity. Most eight-bit microprocessors have buses that function solely in this asynchronous mode. In contrast, some high-performance 32-bit microprocessors may initially boot in a low-speed asynchronous mode and then configure themselves for higher performance operation after retrieving the necessary boot code and initialization data from the EPROM.

4.3 FLASH MEMORY

Flash memory captured the lion's share of the nonvolatile memory market from EPROMs in the 1990s and holds a dominant position as the industry leader to this day. Flash is an enhanced EPROM that can both program and erase electrically without time-consuming exposure to UV light, and it has no need for the associated expensive ceramic and quartz packaging. Flash does cost a small amount more to manufacture than EPROM, but its more flexible use in terms of electronic erasure more than makes up for a small cost differential in the majority of applications. Flash is found in everything from cellular phones to automobiles to desktop computers to solid-state disk drives. It has enabled a whole class of flexible computing platforms that are able to upgrade their software easily and "on the fly" during normal operation. Similar to EPROMs, early flash devices required separate programming voltages. Semiconductor vendors quickly developed single-supply flash devices that made their use easier.

A flash bit structure is very similar to that of an EPROM. Two key differences are an extremely thin dielectric between the floating gate and the silicon substrate and the ability to apply varying bias voltages to the source and control gate. A flash bit is programmed in the same way that an EPROM bit is programmed—by applying a high voltage to the control gate. Flash devices contain internal voltage generators to supply the higher programming voltage so that multiple external voltages are not required. The real difference appears when the bit is erased electrically. A rather complex quantum-mechanical behavior called *Fowler-Nordheim tunneling* is exploited by applying a negative voltage to the control gate and a positive voltage to the MOSFET's source as shown in Fig. 4.5.

The combination of the applied bias voltages and the thin dielectric causes the charge on the floating gate to drain away through the MOSFET's source. Flash devices cannot go through this program/erase cycle indefinitely. Early devices were rated for 100,000 erase cycles. Modern flash chips are often specified up to 1,000,000 erase cycles. One million cycles may sound like a lot, but remember that microprocessors run at tens or hundreds of millions of cycles per second. When a processor is capable of writing millions of memory locations each second, an engineer must be sure that the flash memory is used appropriately and not updated too often so as to maximize its operational life. Products that utilize flash memory generally contain some a management algorithm to ensure that the erasure limit is not reached during the product's expected lifetime. This algorithm can be as simple as performing software updates only several times per year. Alternatively, algorithms can be