

How we made DRAM

The emergence of dynamic random access memory (DRAM) in the 1970s had a huge impact on the future of digital computing. Its inventor, Robert H. Dennard, explains how the drive for simplicity led to this breakthrough.

Robert H. Dennard

I received my PhD from Carnegie Institute of Technology in Pennsylvania in 1958 and joined IBM Research in a period of rapid growth of mainframe computers. After completing several interesting projects, in 1964 I joined a microelectronics group led by Dale Critchlow that was developing a new metal–oxide–semiconductor (MOS) transistor technology and learning how to use it to do computer functions in integrated circuits. The goal was to replace magnetic-core random access memory (RAM), which was large, slow and power-hungry. We wanted to build RAM using integrated circuits with six small MOS transistors in a flip-flop memory cell to store each bit.

A key event in my life occurred on 9 November 1966. I attended a large IBM Research conference and was impressed by a presentation given by Dick Matick, who was part of a group trying to improve magnetic-core memory and keep that technology alive. They had a very simple memory cell, just a small square of thin-film magnetic material between two copper lines on a printed circuit board. I was inspired to find something as simple for the technology we were developing in the MOS group.

I went home that evening to my house in Westchester County, New York, sat down to admire the Croton Gorge view from my living room sofa and continued thinking. I considered the possibilities of using a small, simple, MOS capacitor as a basic memory element, storing a bit of data as an electric charge or no charge (binary 1 or 0). Each memory cell in a two-dimensional memory array would have a transistor connected to a data line to control writing the charge to the capacitor. In my initial thinking, the capacitor was the gate of another transistor, and reading was accomplished by monitoring the current flow in that second transistor. Excited about this idea, I called my boss, Critchlow, even though it was getting late. He listened and then suggested we talk about it tomorrow — he basically told me to take two aspirin and call him in the morning!

When morning came, I recognized that the idea was not as simple as I thought. It required multiple access lines, with

complicated drive schemes or an additional transistor to make a memory array function properly. I kept working on different configurations for a few weeks until a ‘eureka’ moment — I realized that the stored charge could be read back out through the same transistor from which it was written, and it would create a small detectable signal on the data line. The cell had been reduced to a single transistor and a capacitor at the intersection of two access lines (Fig. 1) and thus was much less complex than the six-transistor cell.

My new memory idea was called the single-transistor memory cell because passive components, such as capacitors and resistors, were not counted in those days. One of the characteristics of the single-transistor memory cell is that a small leakage current in the transistor discharges the capacitor in less than a second. (This gives rise to the name ‘dynamic’, as the data bit is stored only temporarily.) To preserve the data, each bit must be refreshed by reading it out and writing it back into each cell at certain intervals. Fortunately, the storage time before the charge leaks off allows many useful memory accesses between the refresh operations.

In 1967, IBM filed for a patent on my single-transistor dynamic random access memory, which became known as DRAM. The patent (Fig. 1) was issued in 1968. In 1970, Intel built the first commercially successful DRAM chip, called the 1103, using a three-transistor memory cell. By the mid-1970s, several manufacturers had introduced 4-kilobit DRAM chips based on the single-transistor memory cell.

The simplicity and low power consumption of DRAM changed the computing industry. It allowed RAM to become very dense and inexpensive. As a result, mainframe computers could be equipped with relatively fast RAM to act as a buffer to the increasing amount of data stored on disk drives. This vastly sped up the process of accessing and using stored information. Just as important, DRAM was combined with microprocessors to make personal computers possible.

At the International Electron Devices Meeting in Washington, DC in 1972,

June 4, 1968 R. H. DENNARD 3,387,286
FIELD-EFFECT TRANSISTOR MEMORY
Filed July 14, 1967 3 Sheets-Sheet 1

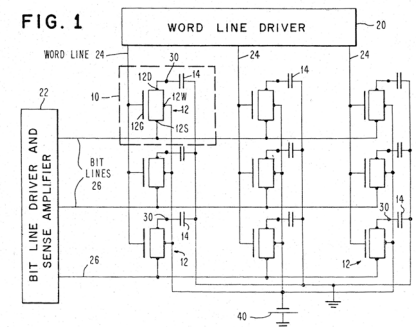


Fig. 1 | Schematic of DRAM from the original 1968 patent. The schematic shows an exemplary small memory array with nine cells. A signal on one of the vertical word lines turns on all the transistor switches (with gates 12G) in that column, connecting all the capacitors (labelled 14) to the lateral bit lines to either ‘write’ data from the bit line drivers or ‘read’ data into the sense amplifiers.

I introduced scaling principles for MOS integrated circuits, which would later be referred to as Dennard scaling. These showed how the transistors and interconnect lines could be made with smaller dimensions, achieving faster circuit speed and much lower power consumption, when the power supply voltage is also scaled down. The idea caught on, and the microelectronics industry strove for continual scaling, leading to the geometric growth in computer performance observed over the past four decades. DRAM was the large-volume predictable product that drove scaling. And, amazingly, today’s DRAM chips contain a million times more bits than those first 4-kilobit chips that launched personal computers. □

Robert H. Dennard
IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA.
e-mail: dennard@us.ibm.com

Published online: 13 June 2018
<https://doi.org/10.1038/s41928-018-0091-3>