

Dreaming Spires

Andrew J. Walker

City skylines tend to rise. Even the emperor Augustus had to place limits on the height of buildings in ancient Rome. While the impetus to rise at that time may have been more to do with status, the main reason nowadays is the ever increasing cost of a square metre. We all need our personal space so there is a limit to how far an apartment or office can be shrunk. By building up rather than encroaching out, the total cost of housing millions of inhabitants and office workers can be minimized. Swap silicon memory cells for buildings and electrons for inhabitants to describe how digital photos, songs and data could be stored in the future.

In the semiconductor industry miniaturization has been the driver of the microelectronics revolution. The approximate doubling of the number of transistors on a silicon chip every 18 months is described by Moore's Law, named after Gordon Moore, one of the founders of Intel Corporation, who predicted this trend over 40 years ago. An obvious consequence is the increasing storage capacities of digital audio players, USB drives, cell phones, digital cameras and increasingly solid state drives for ultra-thin laptop computers. The storage technology used in all of these applications, NAND Flash, was invented by the Japanese company Toshiba over 20 years ago. The basics of this Flash have not changed in all this time. A single cell stores data by placing electrons on a small conducting material that is electrically isolated from anything else on the chip. This storage is nonvolatile in that the electrons remain in place even without electrical power supplied to the chip. The movement of electrons to and from this "floating gate" is governed by a quantum mechanical principle called tunneling that has been known since the 1920's. The presence or absence of these

stored electrons is inferred by their effect on the free flow of electrons in a transistor: the more stored electrons, the less current will flow since electrons repel electrons.

The ability to shrink the size of the floating gate along with its associated transistor has been the reason for such bounds in personal data storage. Since its inception, the physical footprint of one memory cell has shrunk by more than 1600 times. In 1988, the transistor was about 1000 nanometres long. The latest chip from IM Flash Technologies, a joint venture between Intel Corporation and Micron Technology in the USA, has shrunk this to 25 nanometres and contains 32 billion cells all within an area of 167 square millimeters. Through further innovation called Multilevel cell NAND (MLC) each memory cell is able to store 2 data bits instead of just one. This doubles the total storage capacity of the chip. Not to be outdone, two Korean Flash manufacturers, Samsung and Hynix, followed in quick succession to IMFT to say they have NAND Flash technologies waiting in the wings at 27 and 26 nanometres respectively. These numbers define the total number of chips on a silicon wafer. The more chips per wafer, the cheaper each chip can be made. This leads to either greater margins for the manufacturers or greater market share if they can undercut their competition. A nanometer is about 5 silicon atoms placed side by side. Since the total NAND Flash market is predicted to grow towards 50 billion dollars in 4 years, such atomic level differences become a life and death struggle for market share.

Amid all the expected trumpet fanfares announcing these new chips, several notes of caution can now be heard. It has become expected that any new technology will be about 70% linearly shrunk from the previous one. This shrink factor is starting to taper off. The stored electrons in the floating gate are starting to get uppity and complain about their restricted living space. The transistor that is used to sense the presence of electrons in a floating gate is starting to be affected by electrons in other floating gates. It's as if the single shout from one floating gate as heard from its associated transistor is being drowned out by a cacophony of screams from all the floating gates nearby. This electrical

coupling is seen as one of the major challenges to further NAND Flash shrinking. Any closer and the transistor will not be able to make out whether or not there are any electrons in its floating gate.

No major NAND Flash manufacturer can openly admit that they can no longer shrink. Customers like Apple have come to expect a capacity doubling in Flash almost every year with this capacity being passed on to their customers in the form of iPods with greater storage.

Nevertheless, two main technology types as NAND Flash replacement have started to appear in the literature, namely Resistive Random Access Memory (RRAM) and Charge Trap Flash (CTF). What is also new is that they are being developed in a three dimensional stackable approach. Layer after layer of memory cells are being placed one on top of the other to increase chip capacities as an alternative or in addition to lateral shrinking.

In RRAM a material is chosen based on its ability to switch between different nonvolatile resistive states. Each memory cell then contains a small blob of this material with its resistance being programmable and read out with some form of device akin to the read transistor in standard Flash. There are many possibilities for the switching material. In December 2009 at the International Electron Device Meeting in Baltimore, Intel and Numonyx presented a paper on a RRAM which contained a material known as Phase Change Memory based on chalcogenide glass whose switching capabilities have been known for many years. SanDisk, a Silicon Valley memory pioneer, together with Toshiba of Japan are working on some form of RRAM where the material has not been publicly disclosed. Samsung, the Korean company that has the biggest Flash market share, has published a RRAM approach that consists of a simple metal oxide as the switching material. In typical Silicon Valley fashion, the huge market and the impending end of the incumbent technology have spawned several startups looking into RRAM alternatives.

CTF depends on electron storage but instead of a floating gate conductor, the electrons are placed in traps that exist in some form of insulator, usually silicon nitride. This is closer to the existing Flash

approach in that tunneling is used to place the charge in the traps while their presence is sensed with an associated transistor. The troublesome electrical interference from electrons not associated with a particular transistor is much less than in the floating gate case. Covering their bets no doubt, Toshiba is also working on a three dimensional CTF technology. Samsung too has published several CTF technologies in three dimensions. Besides these behemoths, there is a Silicon Valley startup called Schiltron Corporation whose approach solves many of the difficulties associated with implementing CTF in a three dimensional Flash process.

In the notoriously conservative world of silicon manufacturing we may soon see a new approach being adopted where our digital photos and songs will be stored in three dimensional stacks of either stored electrons or as different resistance states in some material new to the chip industry. Instead of leading the way, the microelectronics industry may now be following the ever rising city skylines and for the same reason of packing in the punters at minimum cost.