The non-volatile memory market is growing rapidly with the continuous evolution of Flash technology to smaller geometries delivering higher density and lower cost per bit. At the same time, the convergence of improved intelligence, connectivity and intuitive interfaces is driving new applications and continuous market growth for non-volatile memories.

Non-volatile memory applications can be divided into standalone and embedded system solutions. Standalone applications tend to be driven primarily by costs and the constant need for next generation products. And while cost is also important in embedded solutions, longevity, expanded functionality (extended temperature ranges, higher reliability, higher performance, low power, etc.), and support are required for embedded systems. The embedded market relies mainly on NOR Flash for critical applications such as code storage due to its reliability and performance while NAND is typically utilized for less critical data storage.

In recent years, NAND technology has been scaling more aggressively than NOR and is facing greater challenges in scaling to smaller geometries. Therefore, the industry is evaluating a number of exploratory technologies with the aim of achieving scalability for higher densities and reduced costs at acceptable levels of performance and reliability.

**Hitting the Limits with Floating Gate NAND**
While the industry has found ways to push NAND scalability, it’s looking extremely challenging for the industry to scale at the same pace beyond 20 nm using current Floating Gate technology.

The main challenge in scaling in Floating Gate technology is the reduction in spacing between neighboring cells. The smaller spacing makes it difficult to fill the necessary interlay dielectric and control gates between the neighboring Floating Gate cells. The close proximity of the cells also significantly increases the capacitive coupling causing neighbor cell disturb.

**Charge Trap and Three-Dimensional (3D) NAND**
In order to continue to scale NAND Flash, the industry is evaluating two possible ways to scale below the 20 nm level: 1) Three-Dimensional cell; and 2) planar cell, such as Charge Trap NAND.

3D Flash technology takes a “stacking” approach. One can look at it as if the technology layers are stacked, like a high rise building each floor representative of a different memory cell. The challenge with this technology is that you have very tall “stacks” with a very high aspect ratio (height over width) which creates processing challenges in uniform patterning and depositing materials. In addition, the high aspect ratio of the stacked cell structure makes it less mechanically stable.

Planar Charge Trap NAND technology differs from the more conventional Floating Gate technology...
in that it uses a very thin layer of silicon nitride film, around 100 angstrom (Å), to store electrons rather than the over 1000 Å of a typical Floating Gate storage layer. The thinner layer of silicon nitride significantly reduces the neighboring cell capacitive coupling and eliminates the challenge of filling interlay dielectric in narrow and high aspect ratio spaces between the Floating Gates.

Figure 1. Planar Charge Trap NAND Architecture.

**Title-1**

**Resistive RAM**

Another new non-volatile technology that is being explored by a number of companies, universities and national laboratories is Resistive RAM or RRAM. RRAM relies on two different resistance states to define the zero and one state of the memory. This is achieved by changing the conductivity of the dielectric material between two metal contact layers.

Identifying the right material combination is a key area of RRAM research. Conductive bridge-based and oxygen-based are the more promising materials being explored.

Conductive bridge-based RRAM consists of two metal contact layers (which include material such as silver in one of the layers) with a high-resistance insulator in-between them. A voltage is applied between the metal layers forming a silver-based conductive filament causing a low resistance state. When you change the direction of voltage, the filament breaks between the metal contact layers and creates a high resistance state. These are the two states of the memory.
The biggest challenge with conductive bridge-based RAM is being able to create repeatable, reproducible filaments within a cell and within many cells of the memory array. Another concern is that because conductive bridge-based RRAM usually relies on materials (such as silver) not commonly found in silicon manufacturing facilities, there are contamination and processing challenges introducing these new materials into silicon fabs.

Oxygen-based RRAM uses an oxide insulator between two metal contact layers. An oxide material is chosen and engineered to be non-stoichiometric (it has either an excessive or deficient level of oxygen). When a voltage is applied to the metal contact layers, it creates a path through the insulator that is either high or low resistance, depending on the direction of the bias voltage applied.
Again, the challenge with oxygen-based RRAM is creating a repeatable, reproducible resistance path within the cells and within multiple cells. Oxygen-based material (such as hafnium oxide) is more compatible with existing fabs than conductive bridge-based materials.

To make RRAM a competitive solution requires a very good select device. Some companies are pursuing the use of diodes that are sitting within the metal layers rather than a transistor that is sitting in a silicon substrate so that the geometries are small enough to actually build the next generation of non-volatile memory.

**Phase Change Memory**

Phase change memory also consists of two metal contact layers. In-between the two metal contact layers is a chalcogenide material, mostly based on GeTeSb, which can be in an amorphous insulating phase or polycrystalline conductive phase.

These materials go through phase transformation via a heating and quenching process. The local heating, with temperatures as high as 600 degrees, is generated by flowing current through the material. In the last year or two, a number of companies have introduced low density PCRAM into the market. However, PCRAM has a limited number of program cycles, slow program time, and performs poorly at high temperature. PCRAM also requires high current and high power to achieve the required temperatures for programming, making it difficult to scale.

**Magneto Resistive RAM**

One of the other memories that many companies, universities and research labs are exploring is magneto resistive RAM or MRAM. There are two versions of MRAM, toggle and spin-torque.

MRAM consists of two metal contact layers. In-between the metal contact layers, there are two layers of magnetic material, one fixed in polarization direction and one switchable, separated by an insulator. If the two magnetic materials have the same direction of polarization, the memory cell is in a low resistance state. If they are in polarized in opposite directions, it is in a high resistance.

To program toggle MRAM, a current is applied to the metal lines that are woven below and above the magnetic material, generating a local magnetic field. The sequence of applied currents between the two lines determines the direction of the polarization of one of the magnetic layers in respect to the other.

To program spin-torque MRAM, the current actually flows through the entire material stack. The majority of the electrons within the fixed magnetic layer are spinning in one direction. Depending on the direction of the current passing through the fixed magnetic material, the electrons from the fixed magnetic layer interact with the electrons in the other magnetic layer forcing them to be either in the same or opposite spin direction of the fixed layer.

Toggle MRAM is in production today in densities ranging from 256 Kb to 16 Mb. Toggle MRAM has fast read/write access time, and unlimited read/write cycles. However, scaling toggle MRAM to smaller geometries and larger densities is challenging. Spin torque MRAM is expected to be more scalable while requiring less current to switch. However, spin-torque MRAM is still in research mode. The challenge for spin-torque MRAM is achieving reproducible, repeatable switching while maintaining reliability.

**Moving Forward**

MRAM, PCRAM, and RRAM, in order to become major memory technologies, will have to rely
heavily on new material innovation. On the other hand, Planar CT NAND and 3D NAND leverage existing materials and could fast become commercially viable for high-density Floating Gate NAND replacement.

**About the author**

Dr. Saied Tehrani is senior vice president and chief technology officer at Spansion, responsible for technology development, strategy and new products. He has over 25 years of experience in the semiconductor industry. His work has been recognized with over 75 U.S. patents, over 70 co-authored articles in technical journals, and many awards including the IEEE Daniel Noble award. Prior to joining Spansion, Tehrani was founder and chief operating officer of Everspin Technologies, Inc., served as director and fellow for Analog and Mixed Signal Technologies at Freescale Semiconductor, and held several leadership positions during his 20 years at Motorola. Tehrani holds a doctorate and master's degree in electrical engineering from the University of Florida and a bachelor's degree in electrical engineering from the University of North Carolina.

If you liked this and would like to see a weekly collection of related products and features delivered directly to your inbox, [click here to sign up for the EDN on Systems Design newsletter](#).