Using Hot Carrier Injection for Embedded Non-volatile Memory

- Kenji Noda, Executive VP and CTO, NSCore, Inc.

Since the invention of NAND Flash memory in the 1980's, non-volatile memory (NVM) has touched all aspects of our daily life. Applications for NVM is growing everyday from handy data storage devices like a USB Flash Drive to High Definition Multimedia Interface (HDMI) and its high-bandwidth digital content protection (HDCP) encryption algorithms to firmware storage in a variety of digital devices like digital cameras and mobile phones. Some CPU and ASIC designs still use legacy non-volatile memory technology such as fuse or electric fuse (e-fuse), but the bit cell is larger than newer technologies so it is not so attractive for higher bit count memory arrays and it is typically not field programmable. Other designs use embedded Flash, but current Flash memory technology requires a special fabrication process if it is embedded in existing CMOS logic; this significantly increases the chip cost. Finally, in the last few years, newer NVM technologies such as oxide-rupture and floating gate have entered the market and have been incorporated into more and more designs. These technologies do not have the area penalty of fuse/e-fuse and do not have the additional processing costs of embedded Flash, but have their own unique disadvantages in terms of process scaling, testability, and security. A new technology that uses hot carrier injection for non-volatile data storage is now available as an embedded non-volatile memory solution. Similar to floating gate and oxide-rupture, this technology does not have the limitations of fuse/e-fuse and Flash, but unlike those technologies, it also performs well in terms of process scaling, testability, and security.

Data Storage

The hot-carrier effect has been well known and deeply researched over the last quarter century. Typically in CMOS logic, the desire is to avoid this effect because it changes the performance of transistors over time and ultimately can cause a chip to fail. However, in the technology presented in this white paper, the trapped charge from the hot-carrier effect is used to program data in the bit cells of the NVM.

As shown in Figure 1 below, electrons flowing from the source to the drain in a NMOSFET gain extremely high energy when drifting in a high electric field near the drain edge. Some of these

electrons are injected into insulators such as the gate insulator and the side-wall spacer and are trapped. The charged electrons increase Vt and decrease drain current for the NMOSFET. However, these electrons are very stable even at a high-temperature.

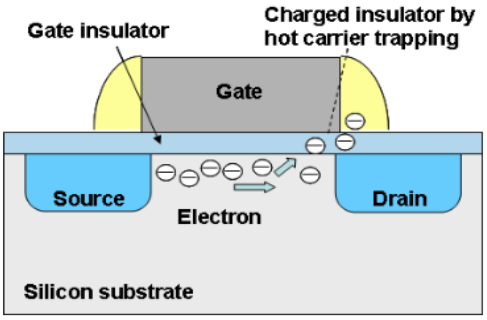

Figure 1. Hot Carrier Trapping

Figure 2 shows Id-Vg curves for transistors as programmed and after baking 200C for 100 hours, as well as showing initial curves as a reference. As demonstrated here and through extensive reliability testing, the transistor, once programmed, has excellent retention characteristics.

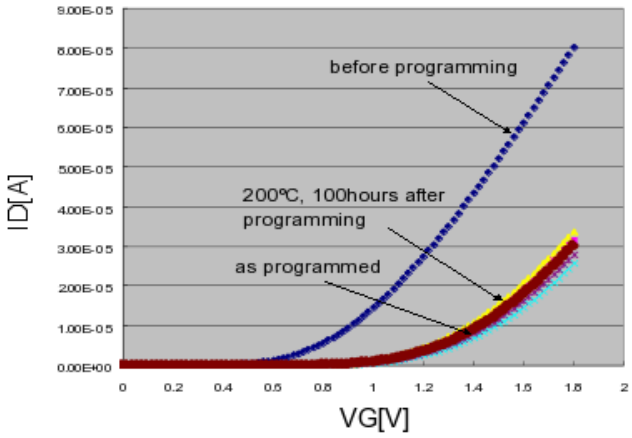

Figure 2. Id-Vg Curve for Unprogrammed and Programmed Bit Cell

Retention characteristics generally depend on distance between the trap site and silicon-insulator interface and barrier height for the site to de-trap. The technology shown here primarily traps electrons in the large volume of insulator at the side-wall spacer and even if there are small pin holes or defects in gate insulator, the retention characteristics are still very

stable. Also, unlike floating gate NVM or SONOS memories that experience additional leakage and reduced lifetime at smaller process geometries or oxide rupture NVM that can experience self-healing in the oxide, the fundamental data storage mechanism described here becomes more robust in smaller process geometries.

Programming the Bit Cell

The schematic diagram shown in Figure 3 shows a bit cell containing the hot-carrier injected transistor highlighted in yellow. The memory bit cell consists of a pair of NMOSFETs, of which the common drain node is coupled to VPP (~5V) for programming operations. The two source nodes are coupled to VDD and VSS, respectively. When the word-line is selected during this programming operation, only the bit cell transistor whose source is at VSS is degraded by hot-carrier effects.

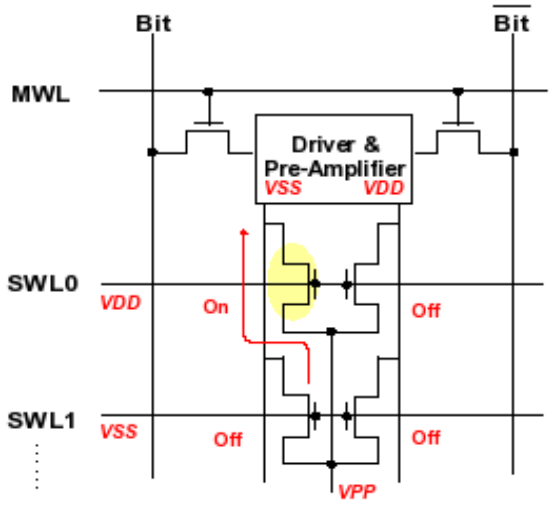

Figure 3. Programming the Bit Cell

Recalling Data from the Bit Cell

To read the stored data from the bit cell as shown in Figure 4, the common node of the bit cell transistors are coupled to VSS and the sense amplifier determines the differential in cell currents due to the shift of the Id-Vg curve of the programmed transistor.

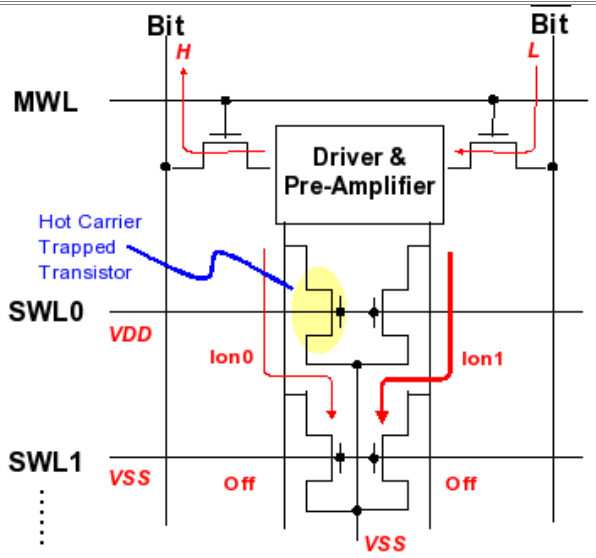

Figure 4. Reading the Bit Cell

By it's nature, the bit cell described here is one-time programmable (OTP), but one can emulate few-time programmable (FTP) or multi-time programmable (MTP) memories with duplicate bit cells and keeping track of the programming count with a store time counter (STC).

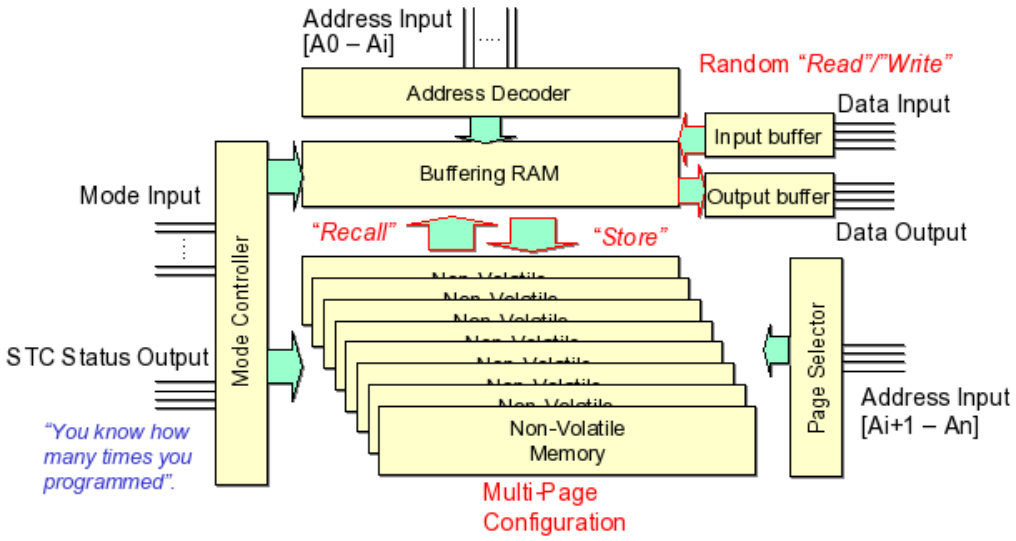

Figure 5. Multi-page NVM Macro Example

Figure 5 below shows a block diagram for an example NVM macro with a multi-page

configuration enabling a high data rate. To recall data from NVM cells to a buffer RAM, it takes from 20ns to 100ns. Accessing the buffer RAM is as fast as high speed SRAM.

Testability

Testability is a primary concern for any OTP NVM because the bit cell can only be programmed once. Ideally, every transistor and node in the bit cell should be visible and testable without actually programming the bit cell. In the bit cell shown in Figure 3, every transistor can be tested for proper 'on' and 'off' current and if there are any connectivity issues in metal wires or vias, that will be detected as well. With this visibility of the transistors and connections, a built-in self test (BIST) can be included in the NVM macro to fully test the memory array as well as other peripheral logic. In the case of oxide-rupture OTP NVM, there may be no visibility to or current flow in the select transistor of the bit cell before programming thus making it impossible to test the select transistor or connectivity in metal wires and vias of the bit cell.

Security

For applications such as encryption key storage for HDMI or even code storage where the code must be protected, not being able to reverse engineer the content of the NVM is critical. With fuse/e-fuse technologies, the memory content can be detected relatively easily through visual inspection. In the hot-carrier charge trapping technology described here, the contents of the bit cell are not visually detectable from the top or side and are not detectable with a FIB voltage contrast.

Summary

With the number of applications for non-volatile memory growing every day, a production proven new hot carrier injection NVM technology is now available to fulfill the common requirements of field programmability, strong testability, scalability to smaller process geometries, and secure data storage.

About the Author

Kenji Noda is responsible for developing NSCore's designs and for technical support to customers.

Prior to joining NSCore, Mr. Noda was Department Manager at NEC Electronics Corporation. There he conducted development on the embedded DRAM process integration and launched a series of new technologies and products. He has a combined experience of 18 years in the areas of SRAM and DRAM design and CMOS process integration from 0.25 to 90-nm generations including embedded DRAM technologies. He has served in technical program committees at CICC for six years. He has authored numerous technical papers on semiconductor device and circuits, especially in the area of memory. He has received BS and MS degrees in Electrical Engineering from Waseda University. You can contact Kenji at noda@nscore.com.

About NSCore

Founded in 2004, NSCore develops, licenses, and markets innovative non-volatile memory technologies for SoC semiconductors which are implemented on high volume, popular CMOS processes without extra steps, masks or process modifications. NSCore's patented PermSRAM® offers the optimum combination of extremely small MTP macro size, fast READ/ WRITE similar to SRAM with very low cost. PermSRAM is not only scalable to 65nm generation and beyond but is also fully testable which is unique to NSCore. PermSRAM's excellent process portability, high yield and automotive level reliability gives SoC design engineers the tools to conceive and build cost effective products with very short TAT. For more information, please visit www.nscore.com.