DIELECTRIC SCALING CHALLENGES AND APPROACHES IN FLOATING GATE NON-VOLATILE MEMORIES

Stephen N. Keeney

Intel Corporation, RN3-01, Santa Clara, CA 95052
Email: stephen.n.keeney@intel.com

ABSTRACT
Dielectric scaling in non-volatile memories (NVM) is approaching the point where new approaches will be required to meet the reliability and performance requirements of future products. For both the tunnel oxide and the inter poly dielectric (IPD), high k materials are being explored as possible candidates to replace the traditional SiO_2 and ONO (Oxide/Nitride/Oxide) films used today. New storage node concepts are also becoming attractive as an alternative approach to address some of the dielectric scaling limitations. This paper will review the current status and discuss the approaches being explored to provide dielectric scaling solutions for future non-volatile memory products.

INTRODUCTION TO FLASH CELL OPERATION

A floating gate non-volatile memory flash cell is shown schematically in Fig.1. It is an MOS transistor with two gates; a floating gate and a control gate. The threshold voltage of the device can be changed by modifying the charge on the floating gate which can retain this charge for many years. Data can be stored in the memory by adding or removing charge. In its simplest form two threshold levels (high and low) can store one bit of information in each memory cell. This concept can be extended to store more levels which is commonly call MLC (Multi Level Cell). Four levels would allow two bits of data per cell to be stored. Adding and removing charge from the floating gate is normally achieved by F-N tunneling or hot carrier injection. A detailed overview of flash cell operation can be found in the publication by Pavan et al. [1]

DIELECTRIC REQUIREMENTS

The tunnel oxide has the most stringent requirements in a flash cell with properties summarized by the following. 1) It must provide a good interface to the silicon channel for reliable transistor operation 2) It should provide efficient charge transport either through tunneling or hot carrier injection during the programming and erase operations which allow the data to be changed in the memory cell 3) It should enable years of retention of the charge on the floating gate. Fig. 2 summarizes the requirements with an Electric Field versus Current plot for a tunnel dielectric. It is desirable to have a large slope to this characteristic to give fast programming and erase at high fields but very low leakage at low fields to provide good retention. It is also important to study the characteristic after stress since SILC (Stress Induced Leakage Current) can be the limiter for retention since it increases leakage current at low fields. The IPD is not expected to transport charge but must block any leakage during the programming and erase operations. It should also deliver years of retention of charge on the floating

gate and provide good capacitive coupling of the control gate to the floating gate to lower the voltages needed for read, program and erase operations.

TUNNEL OXIDE SCALING

SiO_2 thickness limitations have been reported with 60nm being the lower limit due to direct tunneling but a more practical limit of >80nm has been proposed [2] due to post cycling SILC limitations. In actual fact the ITRS tunnel oxide scaling roadmap has been made more conservative in the 2001 edition due to the challenges in going below 80nm. It is becoming clear that to enable a significant reduction below 80nm (EOT) there needs to be a significant change in the approach which until now has focused on SiO_2 and nitridation techniques using NH_3 and N_2O [3]. Several alternatives have been presented in the literature with two of them selected here for further discussion namely JVD (Jet Vapor Deposition) Nitride [4,5,6] and tunneling barrier engineering [7,8].

JVD Nitride

Traditional CVD nitride is a poor choice for a tunnel oxide due to the high trap density leading to Poole-Frenkel conduction (trap assisted transport) resulting in poor retention under low field bias along with poor interface quality. A JVD Nitride has been shown to have low trap density and interface states making it a possible candidate for the flash cell tunnel oxide. This is attributed to the low hydrogen incorporation in the film as a result of the deposition conditions [4]. Data has been reported [5] showing good endurance and reduced SILC (stress induced leakage current) which will improve retention and enable a thinner EOT (Equivalent Oxide Thickness). There is the additional benefit of a lower barrier height for electrons shown in Fig 3. (2.1eV vs 3.1eV for SiO_2) leading to more efficient electron injection resulting in improved performance or scaled voltages. Similarly for holes the barrier is only 1.9eV vs 4.9eV for SiO_2 making hot hole injection a candidate for discharging the floating gate. Although the single cell fundamentals look encouraging there has been little data reported yet on the dielectric performance in a memory array. For this to be a viable replacement for SiO_2 the statistical performance of the film needs to be studied in terms of erase uniformity and data retention on a large sample of cells.

Tunnel Barrier Engineering

Barrier engineering approaches [7,8] are applicable to F-N tunneling program and erase operations. With SiO_2, or any material with a uniform barrier, you can not achieve high transparency at electric fields necessary for fast program and erase operations (~10 MV/cm) along with low transparency at fields (1-3MV/cm) where good retention is required. The reason for this is that the barrier height remains fixed regardless of the applied electric field and field emission occurs due to thinning of the barrier width (Fig.5a). If a triangular or crested barrier approach is used (Fig.5b) then the barrier height reduces under the high field condition when charge transport is desired but increases again when a low field is present reducing the low field leakage current and improving retention. The challenge here is to find the material combination that meets these barrier needs. One possible solution is a trilayer stack of

$Si_3N_4/AlN/Si_3N_4$ which is discussed in reference [7]. It is necessary to provide a very low trap nitride otherwise low level leakage will be high due to P-F conduction. There is a possible opportunity here for a JVD nitride approach but there may be alternative material options that are more attractive. Other such approaches reported by Govoreanu et al. [8] explore high k dielectrics in combination with SiO_2. In this paper both Al_2O_3, Y_2O_3, ZrO_2 and HfO_2 are studied with promising results. Integration of these dielectrics into non-volatile memory devices is underway which will help address the next level of concerns with these approaches.

INTERPOLY DIELECTRIC (IPD)

The key requirements from this dielectric are simple; provide good capacitive coupling to the floating gate from the control gate and minimize any leakage through the dielectric. Unlike the tunnel oxide this dielectric stack is not expected to support charge transfer during programming and erase operations. For most of the last decade ONO scaling has been achieved by thinning the dielectrics while maintaining control of the thickness. This scaling has also been enabled by improving the quality of the oxide on the floating gate which can be aided by reducing the poly surface roughness. Nitridation techniques can also be deployed for the oxide portions of this stack similar to those used to improve the tunnel oxide and this will likely enable scaling into the 100nm EOT thickness [9]. Although it is difficult to predict when this approach will no longer meet the needs due to direct tunneling scaling limitations there would be a clear advantage for voltage scaling if higher capacitive coupling could be achieved. One such approach would be to replace the IPD with a high k material. To maintain the high capacitive coupling improvement a metal gate would be required to reduce the series capacitance from gate depletion. Since this film is not expected to transport charge (as is the case with the tunnel oxide) it will be easier to find a candidate to meet the cell requirements. Replacing the entire ONO stack with a single high k material may be impractical due to the difficulties in depositing a high k film directly on silicon so an SiO_2 interfacial layer may still be required. This bottom oxide will block electron injection but a high k material along with the top gate electrode must be chosen which blocks hole injection. Possible options include Al_2O_3 with a TaN gate which has been proposed for a SONOS type memory device [10] but could in principle be deployed in a floating gate device as well. However, the gate electrode should be chosen consistent with the high k film to make sure the energy barrier is engineered to reduce carrier injection. Going forward much of the development in this area will be engineering the right combination of dielectric and gate materials.

ALTERNATIVE FLOATING NODES

Rather than solving the scaling challenges of the tunnel oxide, alternative floating node concepts are being explored. The basic concept here is to limit the source of electrons rather than trying to maintain the insulation with thinner dielectrics. There has been substantial activity in this area which includes two basic approaches. The first is to replace the polysilicon gate with Silicon Nitride [11] the second looks to use nano-crystals as storage nodes [12,13]. In principle these approaches

can also adopt the IPD scaling options discussed in the previous section.

Nitride Floating Gates

The key concept is that a nitride gate has low conductivity so any tunnel oxide pin hole will not allow discharge of the entire floating gate as would happen with doped semiconductors such as polysilicon. This local storage effect can also be taken advantage of to provide multi bits per cell storage. There are also barrier height changes that effect conduction and retention. Several variations on this theme have been reported in the literature but key differences center around the programming and erasing techniques and the film thicknesses. Fig.6 shows the basic structure of the cell with an ONO film between the gate and the channel. In this case storage occurs in the nitride film above either bitline junctions so two bits of information can be stored in the device which alternates the source and drain to read out the information. Some approaches use hot carrier injection and tunneling while others are focused on tunneling only. In general this approach integrates easily into standard CMOS technology. For those approaches that employ hot carrier injection charge redistribution in the nitride is the primary mode for retention failure (threshold voltage shifts in the flash cell) rather than conduction through the surrounding oxide layers. It is not clear at this time if these limitations can be addressed sufficiently to make this a mainstream alternative to floating gate memories.

Nano-Crystal Storage Nodes

This approach replaces the floating gate with a sea of nanocrystals embedded in an insulator as shown in Fig. 7. Several types of crystals have been explored with most of the literature focused on silicon or metal structures [12] [14] with a good overview published in reference [13]. The key challenge with this approach is to create small, uniform, high density nano-crystal films. These requirements are important for the following reasons 1) Need to have a large number of nano crystals per device so statistical fluctuations from device to device are small 2) Need small uniform size crystals so that we get equivalent electron storage per node and can rely on Coulomb repulsion to self saturate the charge stored on each node. However, it has been observed that as we take this approach to the limit (very small, very dense nano-crystals) they begin to behave like a single floating gate since we can get charge transfer between the crystals so this would represent the scaling limit for this approach [15]. Most development focus is centered on the nano-crystal growth techniques and retention.

Floating Gate To Floating Gate Isolation

So far we have discussed dielectric scaling with regards to isolation of the floating gate from the channel or control gate. One other key concern for NVM scaling is the isolation of floating gates from each other. The main concern is not charge transport between floating gates but capacitive coupling (Fig.8). As device dimensions shrink this can become a limitation since coupling between the floating gate structures can cause a Vt shift on one device when the charge on the neighboring device is changed (as would occur when this cell would be programmed). Both storage node approaches previously discussed to address tunnel oxide scaling will

potentially address this concern since coupling is reduced when thin layers of nitride or nano-crystals are deployed [16]. However, with the high k approaches where the polysilicon floating gate remains, this coupling can limit some of the scaling opportunities and needs to be carefully accounted for during the cell design.

SUMMARY

The industry is facing some significant challenges in the area of dielectric scaling for non-volatile memories. Several approaches for both tunnel oxide scaling and interpoly dielectric scaling have been reported along with alternative flash cell structures that reduce some of the traditional limitations. Many challenges remain before any of these approaches are ready for production and these solutions will compete with alternative memory storage approaches being investigated.

ACKNOWLEDGEMENTS

I would like to thank my Intel colleagues Al Fazio and Stafan Lai for their inputs to this paper.

REFERENCES

[1] Flash memory cells-an overview *Pavan, P.; Bez, R.; Olivo, P.; Zanoni, E.;* Proceedings of the IEEE , Volume: 85 , Issue: 8 , Aug. 1997. Pages:1248 – 1271

[2] Tunnel oxide and ETOX^{TM} flash scaling limitation. *Lai, S.;* Nonvolatile Memory Technology Conference, 1998. 1998 Proceedings. Seventh Biennial IEEE , 22-24 June 1998, Pages:6 – 7

[3] Scaling down of tunnel oxynitride in NAND flash memory: oxynitride selection and reliabilities. *Jonghan Kim; Jung Dal Choi; Wang Chul Shin; Dong Jun Kim; Hong Soo Kim; Kyong Moo Mang; Sung Tae Ahn; Oh Hyun Kwon;* Reliability Physics Symposium, 1997. 35th Annual Proceedings., IEEE International , 8-10 April 1997. Pages:12 – 16

[4] High-quality MNS capacitors prepared by jet vapor deposition at room temperature. *Wang, D.; Ma, T.-P.; Golz, J.W.; Halpern, B.L.; Schmitt, J.J.;* Electron Device Letters, IEEE , Volume: 13 , Issue: 9 , Sept. 1992. Pages:482 – 484

[5] Low-voltage, fast-programming P-channel flash memory with JVD tunneling nitride. *Min She; Tsu-Jae King; Chenming Hu; Wenjuan Zhu; Zhijiong Luo; Jin-Ping Han; Tso-Ping Ma;* Semiconductor Device Research Symposium, 2001 International , 5-7 Dec. 2001. Pages:641 – 644

[6] Extending gate dielectric scaling limit by use of nitride or oxynitride *Wang, X.W.; Shi, Y.; Ma, T.P.; Cui, G.J.; Tamagawa, T.; Golz, J.W.; Halpen, B.L.; Schmitt, J.J.;* VLSI Technology, 1995. Digest of Technical Papers. 1995 Symposium on , 6-8 June 1995 Pages:109 – 110

[7] Riding the crest of a new wave in memory [NOVORAM]. *Likharev, K.K.;* Circuits and Devices Magazine, IEEE , Volume: 16 , Issue: 4 , July 2000 Pages:16 – 21

[8] VARIOT: a novel multilayer tunnel barrier concept for low-voltage nonvolatile memory devices *Govoreanu, B.; Blomme, P.; Rosmeulen, M.; Van Houdt, J.; De Meyer, K.;*

Electron Device Letters, IEEE , Volume: 24 , Issue: 2 , Feb. 2003 Pages:99 – 101

[9] Thickness scaling limitation factors of ONO interpoly dielectric for nonvolatile memory devices
Mori, S.; Araki, Y.Y.; Sato, M.; Meguro, H.; Tsunoda, H.; Kamiya, E.; Yoshikawa, K.; Arai, N.; Sakagami, E.;
Electron Devices, IEEE Transactions on , Volume: 43 , Issue: 1 , Jan. 1996 Pages:47 – 53

[10] A novel SONOS structure of SiO/sub 2//SiN/Al/sub 2/O/sub 3/ with TaN metal gate for multi-giga bit flash memories. *Chang Hyun Lee; Kyung In Choi; Myoung Kwan Cho; Yun Heub Song; Kyu Charn Park; Kinam Kim;* Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International , 8-10 Dec. 2003 Pages:26.5.1 - 26.5.4

[11] NROM: A novel localized trapping, 2-bit nonvolatile memory cell
Eitan, B.; Pavan, P.; Bloom, I.; Aloni, E.; Frommer, A.; Finzi, D.; Electron Device Letters, IEEE , Volume: 21 , Issue: 11 , Nov. 2000 Pages:543 – 545

[12] Fast and long retention-time nano-crystal memory. *Hanafi, H.I.; Tiwari, S.; Khan, I.;* Electron Devices, IEEE Transactions on , Volume: 43 , Issue: 9 , Sept. 1996 Pages:1553 – 1558

[13] Nanocrystal nonvolatile memory devices. *De Blauwe, J.;* Nanotechnology, IEEE Transactions on , Volume: 1 , Issue: 1 , March 2002. Pages:72 - 77

[14] New non-volatile memory with extremely high density metal nano-dots
Takata, M.; Kondoh, S.; Sakaguchi, T.; Choi, H.; Shim, J.-C.; Kurino, H.; Koyanagi, M.; Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International , 8-10 Dec. 2003. Pages:22.5.1 - 22.5.4

[15] How far will silicon nanocrystals push the scaling limits of NVMs technologies? *De Salvo, B.; Gerardi, C.; Lombardo, S.; Baron, T.; Perniola, L.; Mariolle, D.; Mur, P.; Toffoli, A.; Gely, M.; Semeria, M.N.; Deleonibus, S.; Ammendola, G.; Ancarani, V.; Melanotte, M.; Bez, R.; Baldi, L.; Corso, D.; Crupi, I.; Puglisi, R.A.; Nicotra, G.; Rimini, E.; Mazen, F.; Ghibaudo, G.; Pananakakis, G.; Compagnoni, C.M.; Ielmini, D.; Lacaita, A.; Spinelli, A.; Wan, Y.M.; van der Jeugd, K.;* Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International , 8-10 Dec. 2003 Pages:26.1.1 - 26.1.4

[16] High speed and nonvolatile Si nanocrystal memory for scaled flash technology using highly field-sensitive tunnel barrier. *Seung Jae Baik; Siyoung Choi; U-In Chung; Joo Tae Moon;* Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International , 8-10 Dec. 2003 Pages:22.3.1 - 22.3.4

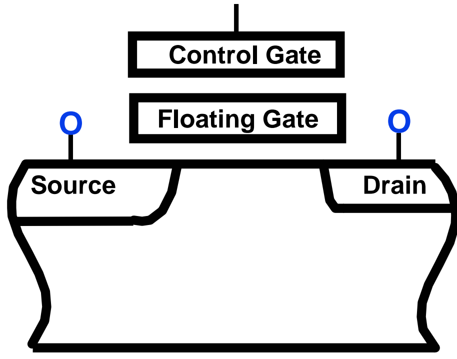

Fig. 1 Schematic picture of a floating gate flash memory device.

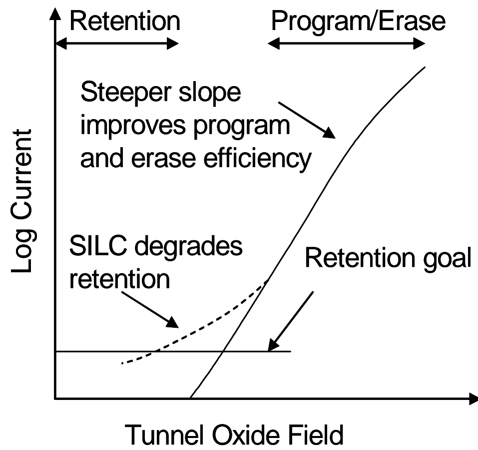

Fig.2 Electric Field versus Current characteristic of a tunnel dielectric showing the dielectric requirements.

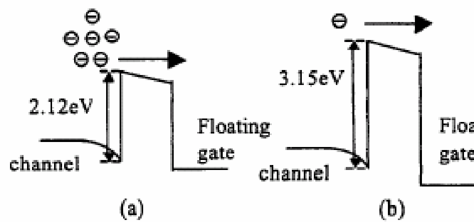

Fig. 3. Lower electron barrier for a JVD nitride tunnel dielectric (a) compared to SiO_2 shown in (b) from ref. [5].

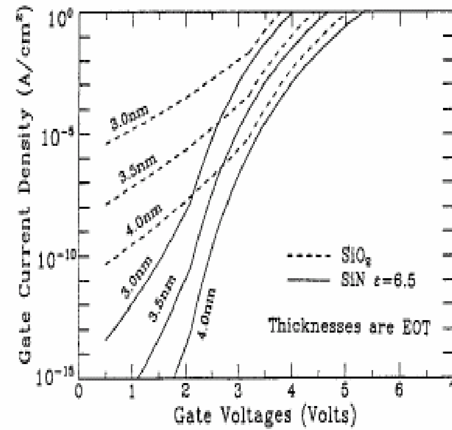

Fig. 4. Dielectric tunneling characteristic showing the *theoretical* improvement in low field leakage with the higher dielectric constant nitride film [6].

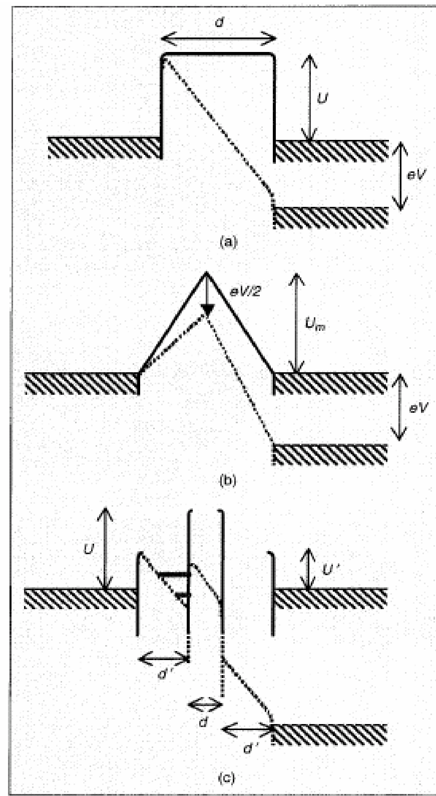

Fig. 5. Classical single barrier (a) shows tunneling is modulated by the barrier width. With barrier engineering it is possible to have a lower barrier under high field and a higher barrier under low field (b) and (c) [Ref. 7]

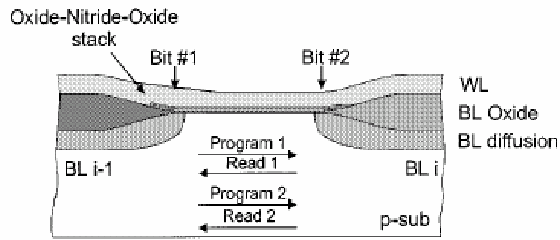

Fig. 6. Nitride storage layer replaces the floating gate allowing tunnel oxide scaling from reference [11].

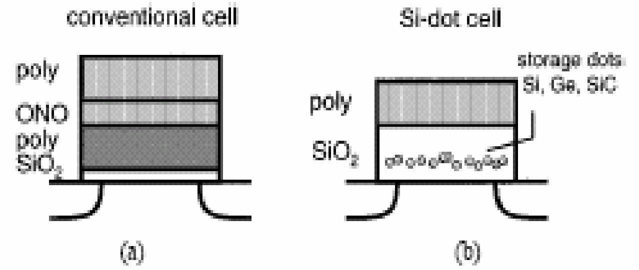

Fig.7. Conventional floating gate cell (a) and silicon nanocrystal memory cell (b) from reference [13].

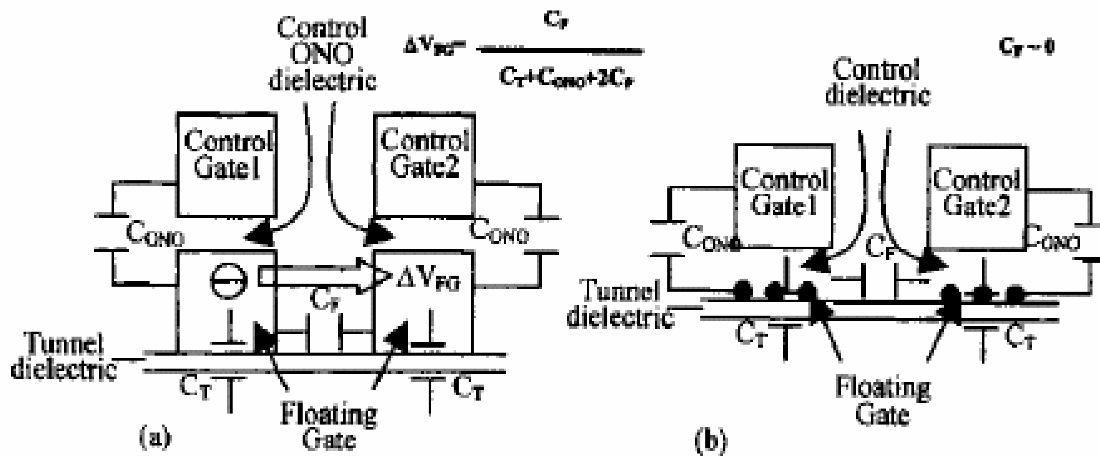

Fig.8. From reference [16] this figure shows the capacitive coupling (C_F) that occurs between the floating gates in diagram (a). Using nanocrystal memory this coupling (C_F) can be substantially reduced (b).