Comprehensive Scaling Study on 3D Cross-Point PCM toward 1Znm Node for SCM Applications

W.C. Chien[1], H.Y. Ho[2], C.W. Yeh[1], C.H. Yang[1], H.Y. Cheng[1], W. Kim[3], I.T. Kuo[1], L.M. Gignac[3], E.K. Lai[1], N. Gong[3], Y.C. Chou[2], C.W. Cheng[3], Y.F. Lin[2], J.M. Papalia[3], F. Carta[3], A. Ray[3], R.L. Bruce[3], M. BrightSky[3], and H.L. Lung[1]

IBM/Macronix Phase Change Memory Joint Project

Macronix International Co., Ltd., [1]Emerging Central Lab, and [2]Product Design and Engineering Center, Hsinchu, Taiwan

[3]IBM T.J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, 10598, USA

TEL: +1-914-945-3575, Email: wcchien@mxic.com.tw; chien@us.ibm.com

Abstract

We present a scaling study toward 1Znm node 3D Cross-point PCM (XPCM) for Storage Class Memory (SCM) applications. The low operation current, and low metal line loading resistance are desired to avoid a wide operation voltage distribution in a cross-point array. For the first time, AC threshold voltage (Vth) of 1S1R OTS-PCM was studied, which will impact the operation scheme. To achieve Tera bits per chip density, six layers 1Znm 3D XPCM with OTS showing high Vth and low leakage current, and scalable periphery circuit are required.

I. Introduction

3D XPCM array using Ovonic Threshold Switch (OTS) with PCM memory is necessary to achieve the cost and capacity requirements for SCM used for either server grade SSD or DRAM-like applications [1]. In this paper, we present the following: (a) a scaling study considering loading effects and operation current; (b) an optimized design with considerations of the Vth of OTS with AC pulses, resistance drift and temperature dependence; (c) the Vth design criteria under "1/2V" [2] and "1/3V" operation schemes [3] to achieve high density OTS-PCM cross-point array; (d) leakage current requirements of the OTS; (e) cost evaluation based on 3D structure efficiency and scaling study toward 1Znm to achieve a competitive Tb level 3D XPCM memory; and (f) a high-density novel information storage using OTS is proposed to record the address of the defects, trimming information, and initial setting information [4] to handle complicated ECC and redundancy coding for 1Znm node. This work points out a roadmap of realizing 1Tb SCM using 3D XPCM for SCM.

II. Scaling Study

Figure 1 shows the IV curves of an 80nm 1S1R OTS-PCM cell using TeAsGeSiSe-based OTS (OTS A) and doped Ge2Sb2Te5 PCM [5]. Two Vths are observed for RESET (VtR) and SET (VtS) states, respectively. The Vth window is determined by the difference between the VtR and VtS. The RESET current of 80nm OTS-PCM is about 700uA [5]. Program voltage (Vpro) distribution is simulated with abovementioned device considering word-line (WL)/bit-line (BL) loading resistance (R_{LOAD}) of 1kΩ as shown in **Fig. 2**. It shows wide Vpro distribution from 4.2V to 5.6V along the best corner to worst corner caused by the operation current of PCM (I_{OPER}) and R_{LOAD} (Vpro_total=Vpro+I_{OPER}*R_{LOAD}). Since the operation current of PCM is reduced by scaling [5], the scaling effect with fixed R_{LOAD} 1kΩ is estimated in **Fig. 3**. The Vpro distribution is tightened when the PCM is smaller than 40nm due to lower operation current. However, the BEOL metal resistivity increases significantly beyond 20nm node [6]. At an extreme R_{LOAD} of 20kΩ, simulation of 14nm device shows wide Vpro distribution and very high Vpro causing "1/2V" scheme to fail (½Vpro>VtS) as shown in **Fig. 4**.

III. Threshold Voltage Design

The AC Vth tests of the VtR using 5ns and 100ns pulses are shown in **Fig. 5**. At 5ns pulse the VtR is about 4.1V, which is higher than DC measurement. Once the pulse width increases to 100ns, the VtR is about 3.7V matching the DC result as shown in **Fig. 1**. Subsequently, **Fig. 6** shows the pulse width dependence of VtS and VtR and we find the maximum Vth window beyond 100ns. To maintain Vth window, additional factors needing consideration is VtS drift (**Fig. 7**), which increases VtS by 0.5V after 10hrs, and temperature (**Fig. 8**), where both forming voltage (Vf) and Vth decrease at higher temperatures. The key factors for optimal Vth window design are shown in **Fig. 9**. High Vth OTS is necessary to fulfill the "1/2V" scheme criteria for both read voltage (½Vread<VtS) and program voltage (½Vpro<VtS, and Vpro>VtR) conditions (black region). In contrast, low Vth OTS can only be used by "1/3V" scheme for programming (yellow region). **Figure 10** shows the VtS and VtR distributions for OTS A with thin PCM and AsSeGe OTS (OTS B) [7] with thick PCM (Vth window is tunable by PCM thickness). OTS A can fit "1/3V" Vpro and "1/2V" Vread schemes while OTS B can fit the "1/2V" Vpro and Vread schemes but high voltage operation is required [8] (**Fig. 9**).

IV. Cross-Point Array and 3D Cost

Figure 11 shows the IV curves of OTS A to investigate the leakage current impact on the cross-point array at different temperatures. A linear relationship was found by plotting $\ln(J/T^2)$ versus $V^{1/2}$, indicating the conduction model may fit the Trap limited conduction model [9] with corresponding barrier height of about 0.35eV to 0.44eV (**Fig. 12**). **Figure 13** shows the leakage current criterion estimation. To maintain reasonable read current of 10uA, less than nA leakage current at 0.7Vth is needed for 1k by 1k cross-point array. The power consumption calculation for OTS A (low Vth device) and OTS B (high Vth device) shows that OTS A demonstrates good read performance but poor write performance which is suitable for the SCM with heavy read function, and OTS B demonstrates both good read and write performances but the high voltage operation is needed which is suitable for the SCM with heavy write function as depicted in **Fig. 14**.

Figure 15 shows the structure of the 3D XPCM with six memory layers. Each memory layer shares either a WL or BL with the adjacent layer. The WL(BL) decoders and sensing amplifier (SA) are under the array. The decoder area is a function of the memory layers. To consider the array efficiency, six layers show the lowest cost because the lowest ratio of decoder area to chip density as shown in **Fig. 16.** To achieve 1Tb single chip, 1Znm 3D XPCM using six layers and scalable periphery circuit are needed as shown in **Fig. 17**.

V. Novel Information Storage Scheme

To handle complicated ECC, and redundancy for high density 3D XPCM, a novel information storage is proposed in **Fig. 18**. The initial state without forming process (non-forming state) is used to store the state '0', while the device after forming associated with SET operation is used to store the state '1'. (Note: The device pre-screen for the initial state can be done by checking the leakage current of the OTS). Following, the memory cell with state '0' by using forming process associating with RESET operation is activated. The new state '0' will be defined by the PCM RESET state. **Figure 19** demonstrates good distributions of the Vf and Vth using OTS A with PCM device after soldering reflow, indicating this novel storage scheme is practical.

VI. Summary

A comprehensive scaling study on OTS-PCM cross-point memory toward 1Z node including 3D discussion is demonstrated. Additionally, a novel information storage scheme is presented.

References

[1] M.J. Kang, et al., IEDM, 27.5, 2018. [2] Yi-Chou Chen, et al., IEDM, 37.4, 2003. [3] Jiantao, et al., IEEE Transaction on Electron Device, p.1369, v. 61, 2014. [4] S. Takase, et al., IEEE J. Solid-Sate Circuits, p. 1600, v. 34, 1999. [5] C.W. Yeh, et al., VLSIT, p. 205, 2018. [6] M. Nail, IEDM, 5.6, 2018. [7] H.Y. Cheng, et al., IEDM, 37.3, 2018. [8] W.C. Chien, et al., IRPS, 2019 accepted. [9] D. Ielmini, Phys. Rev. B 78, 035308, 2008.

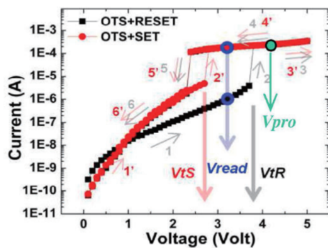
Fig.1 I-V characteristic for 80nm 1S1R OTS-PCM in RESET state and SET state. The Vth of RESET state (VtR) is higher than SET state (VtS). (reprinted from [3])

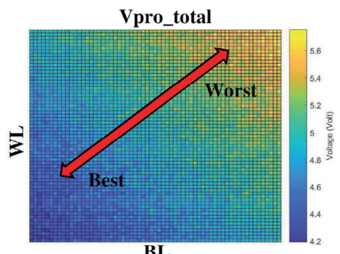
Fig.2 Program voltage (Vpro_total) simulation including loading effect of a cross-point array using 80nm device. WL (BL) loading resistance is 1kohm.

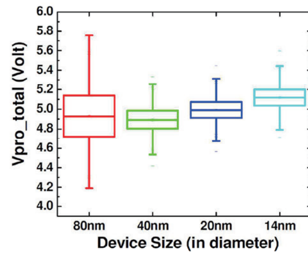
Fig.3 Program voltage (Vpro_total) simulation for different technology nodes. WL (BL) loading resistance is fixed as 1kohm.

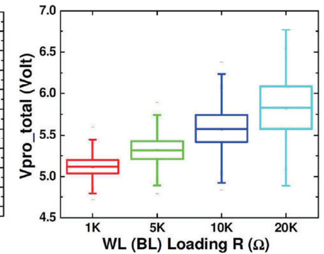
Fig.4 Program voltage (Vpro_total) simulation at 14nm node with various loading resistances.

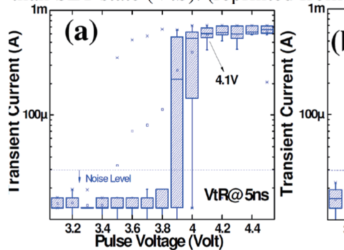
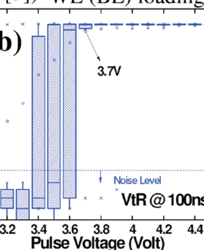
Fig.5 AC Vth test for RESET state using 5ns pulse as shown in (a) and 100ns pulse as shown in (b). Large transient current means OTS turning on indicating the corresponding Vth. VtR is 4.1V at 5ns pulse and 3.7V at 100ns pulse.

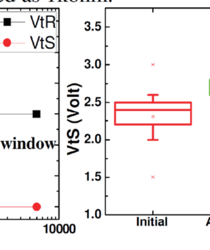
Fig.6 Pulse width dependence to the VtS and VtR. Maximum Vth window occurs beyond 100ns pulse.

Fig.7 The VtS drift test after 10hrs. VtS measurement using 100ns pulse shows drift behavior with the time.

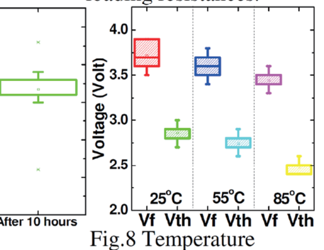
Fig.8 Temperature dependence to the Vf and Vth of 1S1R OTS-PCM. Both Vf and Vth reduce at higher temperatures.

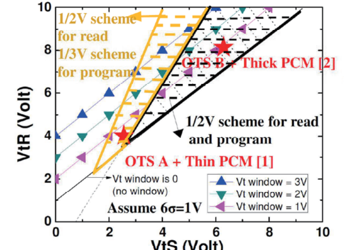
Fig.9 The Vth window calculation. High Vth OTS is required to have enough Vth window, and Vpro window.

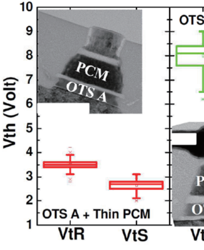
Fig.10 The VtR and VtS distributions for low Vth OTS and high Vth OTS.

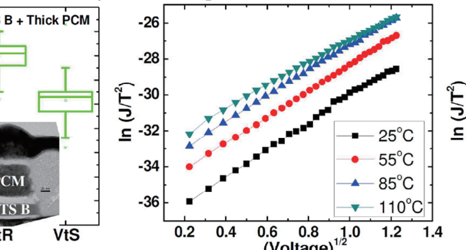
Fig.11 IV curves with various temperatures. The IV fitting matches trap-limited conduction model.

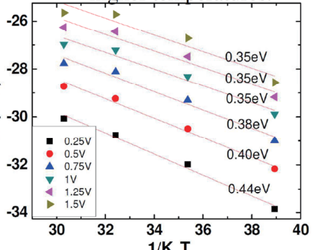
Fig. 12 The barrier height calculation. The barrier height is ranging from 0.35eV to 0.44eV.

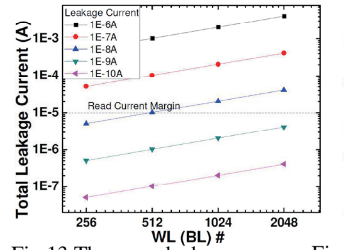
Fig.13 The array leakage current calculation. Device leakage current with nA level is required for 1k by 1k array.

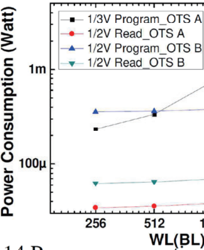
Fig.14 Power consumption estimation by using OTS A (1/3V program, 1/2V read) and OTS B (1/2V program, 1/2V read).

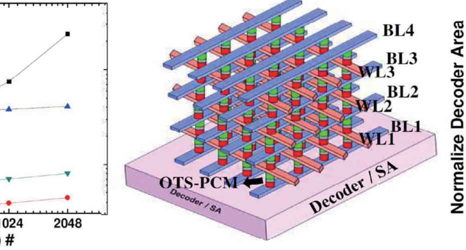
Fig.15 Six layers 3D XPCM. Each memory layer shares with WL or BL. The decoder and sensing amplifier are designed underneath the memory array.

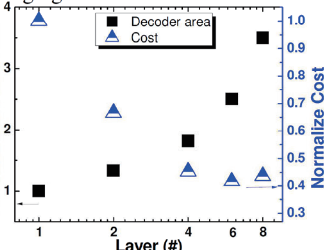
Fig.16 Normalize decoder area and cost versus memory layers. Six layers 3D XPCM shows lowest cost.

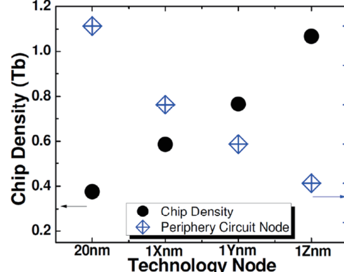
Fig.17 One Tb single chip can be achieved using 1Z node with 6 layers stack when the Periphery circuit node reduces 40%.

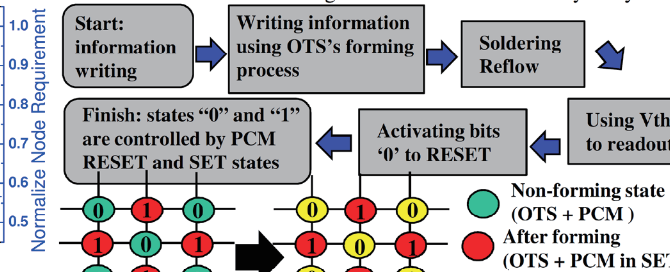

Fig.18 The information are stored in PCM SET state '1' and non-forming state of the OTS '0. Following, the activation mode is used to form the selector and reprogram the state '0' to PCM RESET state (The alternative way is moving the information data to another memory block).

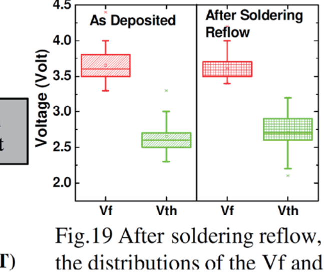
Fig.19 After soldering reflow, the distributions of the Vf and Vth are separated well.