Memory Technology: Process and Cell Architecture

Jeongdong Choe*
TechInsights, 1891 Robertson Road, Suite 500, Ottawa, ON, K2H 5B7 Canada

ABSTRACT

DRAM cell scaling down to the 14 nm design rule (D/R) has already been productized by major DRAM players such as Samsung, Micron, and SK Hynix. They're developing n+1 (12~13 nm) and n+2 (11 nm or beyond) so-called D1b (D1β), D1c (D1γ), and D1d (D1δ) or even D0a generation now, which means DRAM cell D/R might be able to further scale down to single digit nm with EUVL adoption for DRAM cell/core patterning. The cell design scaling down is getting slower due to many scaling issues including patterning, leakage, and sensing margin. Major DRAM players have applied EUVL masks (such as SS_BLP/H_SC2) on DRAM and will expand it for the next generation. Current $6F^2$ cell architecture with 1T+1C will be moving over to $4F^2$ or 3D DRAM in an 8~9 nm D/R DRAM generation due to scaling limitations, which will be D0b or D0c. Although they change to 3D DRAM with a little relaxed CDs, EUVL will be a must for the performance and yield improvement (defects) in DRAM core areas such as very dense WLD and SA patterns just near the cell array. HKMG DRAM process has been adopted on Graphic DRAM and advanced DDR5 DRAM products by Samsung and Micron, although etching and high-k engineering are different for each. Major NAND manufacturers are still in the race to increase the number of vertical 3D NAND gates, they all have already introduced their own 176L/232L/238L 3D NAND devices. Samsung V-NAND, KIOXIA/WDC BiCS, Intel FG CuA, Micron CTF CuA, SK Hynix 4D PUC, and YMTC Xtacking 3D NAND products are the mainstream for SSD and mobile storage applications. Many innovative processes and designs have been adopted, however, lots of challenges are still there to overcome. Although lithography burdens were reduced by changing 2D to 3D, instead, UHAR etching/cleaning/filling-related developments are ongoing. Micron already exited XPoint memory, and Intel is winding it down as well. For SCM applications, fast NAND and some Emerging Memory (EM) devices such as Z-NAND, XL-FLASH, and STT-MRAM will cover the market needs in the future. Due to the difficulties of the EM materials etching, they're currently limited to embedded and low-density applications only. We'll discuss current and future challenges on DRAM, 3D NAND, and Emerging memory including process, design, and materials.

Keywords: Memory, Process, Cell, Architecture, 3D NAND, DRAM, Emerging Memory, Embedded Memory

1. INTRODUCTION

Historically, DRAM cell transistors have been changed from planar to RCAT, and then BCAT to suppress the short channel effect and reduce leakage current. Although VCAT or 3D DRAM cell architecture has not been commercialized yet, the cell scaling is continued by using multi-MESH capacitor integration such as 3- or 4-MESH or more to increase cell capacitance, EUVL or high-NA EUVL, more conformal ALD, and no-plasma-damaged ALE tools. In parallel, $4F^2$ cell design, 3D DRAM, capacitorless DRAM, or adopting the GAA concept to cell structure IGZO-based cell structure may be developed and adopted on 3D DRAM cell integration. For higher DRAM density and bandwidth, more advanced HBM TSV for HBM4 or HBM5 will be needed, and a more compact chiplet package and defect-free hybrid-bonding technology as well. Especially, 3D chiplet can contain high-density memory stacked with the processor in the form of 3D stacking packages. In addition, advanced processing-in-memory types such as HBM-PIM from Samsung or GDDR6-AiM from SK Hynix would be a good direction to accelerate large-scale processing in data centers, high-performance computing systems, and AI-enabled mobile applications in the near future. eSRAM, eFLASH, and some of the embedded and emerging memory devices are currently used for the cache memories. Between DRAM and NAND Flash or SSD storage, there is a speed gap of around 1μs, some storage class memory devices, SCM so-called, such as XPoint or fast NAND such as Z-NAND from Samsung and XL-FLASH from KIOXIA/WD are placed for the SCM. So, emerging memory technology targets eDRAM or eFLASH replacement for cache applications and can be one of the SCM devices as well. We need low power consumption for mobile and wearable applications and data centers, and memory performance is increasingly limiting system performance.

*jchoe@techinsights.com; phone 1-613-404-2797

Optical and EUV Nanolithography XXXVI, edited by Anna Lio, Proc. of SPIE
Vol. 12494, 1249402 · © 2023 SPIE · 0277-786X · doi: 10.1117/12.2658765

eDRAM can be replaced by eSRAM from 10nm or sub-10nm node. eFLASH will be replaced by eMRAM from 22nm technology node. XPoint memory is winding down as Intel announced. Z-NAND from Samsung seems to be on hold now. In the near future, the processing in memory (PIM) application will be placed between the L4 cache and HBM or DRAM products, while compute-express-link (CXL) memories are placed onto the higher-speed SCM. Near data processing (NDP), and in-storage processing applications will be needed for SCM, too. Regarding 3D NAND, each memory company announced or commented on the future 3D NAND scaling and roadmap including up to or more than 800 layers or even 1,000 layers. For further 3D NAND scaling for higher performance and higher density, X-, Y- and Z-direction scaling is a must together with increasing the stacks and peripheral scaling corresponding to higher array density in the near future. For the long-term goal, wafer bonding technology such as hybrid bonded BV-NAND (Bonded V-NAND) and die stacking or logically increasing bit cell level might be needed [1-5].

2. 3D NAND CELL PROCESS AND CELL ARCHITECTURE

To date, there are eight different 3D NAND architectures we've seen in the commercial market. Those are Samsung V-NAND single-deck and recently applied double-deck COP V-NAND, KIOXIA/WD BiCS, Micron CTF CuA, Intel FG CuA, SK Hynix PBiCS, SK Hynix 4D PUC, and YMTC H-bonded Xtacking structure. Among them, Samsung COP-VNAND, Micron/Intel's CuA, SK Hynix 4D PUC, and YMTC Xtacking use the same concept with peripheral CMOS circuitry under the NAND cell array although the trademark names are different from each other. Samsung keeps a single deck up to V6 128L, however, they changed it to a multi-deck and COP concept for the latest 176L. Here, COP means Cell on Periphery, the same concept as CuA and 4D PUC. Intel uses a three-deck cell array structure on CMOS circuitry for the first time for their 144L. YMTC Xtacking uses Hybrid bonding (H-bonding) technology, with two wafer processes, one for NAND array and another for peripheral logic circuitry, followed by through-silicon contact and backend metal layer process. YMTC updated the technology and added some innovative processes such as new materials and new source connection schemes for Xtacking2.0 and 3.0 generations. Unit cell volume is important for cell design, process, device integration, and scaling as well, because it includes one more factor to consider, vertical gate pitch. Unit cell volume scaled down on and on, generation by generation, from 3D NAND manufacturers. Samsung, KIOXIA, Micron, SK Hynix, and YMTC scaled down the unit cell volume for their new CTF COP, CuA, 4D PUC, and Xtacking structure up to 176L and 232L. Intel keeps the same unit cell volume for their recent 144L. Cell volume scaling is one of the big challenges because of the device performance degradation due to cell interference, crosstalk, and process difficulties (Figure 1).

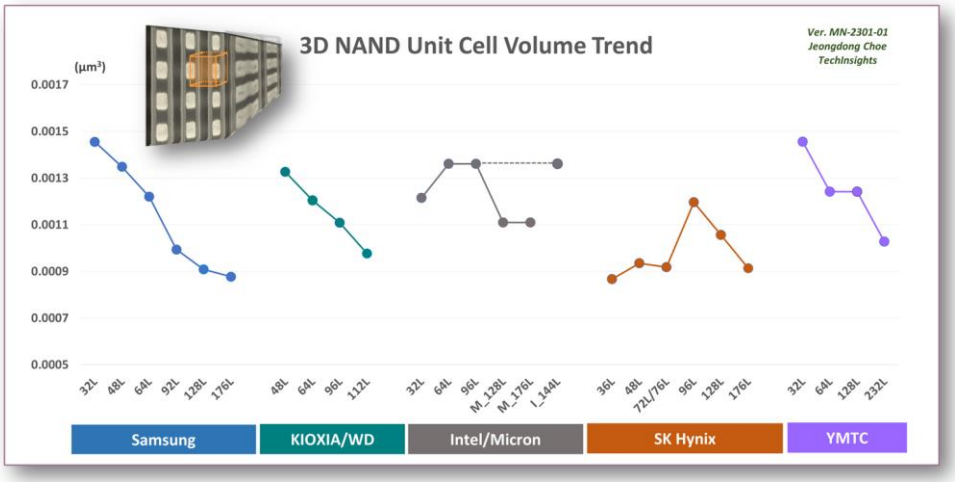

Figure 1. 3D NAND unit cell volume trend up to 232L: Samsung, KIOXIA/WD, Micron, Intel, SK Hynix, YMTC

All the 3D NAND manufacturers adopted 2-decks. It's common now to use a multi-deck cell array structure. Intel already adopted a 3-deck structure. Here, we can compare the number of assigned gates for the upper deck and the lower deck. We may say that they have the same number assigned for each like 50% and 50%, however, they are different by generation, and also different by manufacturer. For example, Micron and Intel's 64L and 96L have exactly the same number of gates assigned for upper and lower decks. Many of the devices have similar numbers for each deck, but Samsung's and SK Hynix's are not that way. Samsung 176L has just 90 + 101, SK Hynix 176L has 90 + 106. Even their

differences have increased. That is due to the number of dummy WLs, passing WLs, and selectors being assigned differently. It'll be good to make the same number for each deck from a process integration viewpoint (Figure 2).

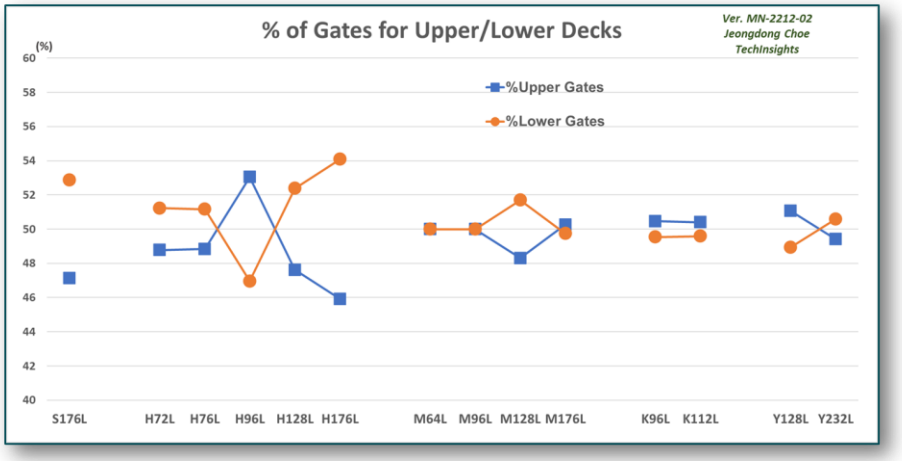

Figure 2. Comparison of 3D NAND cell design: % of gates for upper and lower decks

DRAM cell size has scaled down on and on (Figure 4). The blue square shows Samsung, and the gray circle shows Micron's DRAM cell size. Until the D1x generation between Samsung and Micron, they had a little technology gap, always Samsung had led Micron for each generation. However, for the D1y generation, the gap had narrowed, and all the big three DRAM companies showed the same cell size for the D1z generation. Recently, Micron successfully caught up with Samsung for the D1z generation. Further, Micron released their volume products with D1α generation. Its cell size is measured with a 15% size reduction from the previous D1z generation.

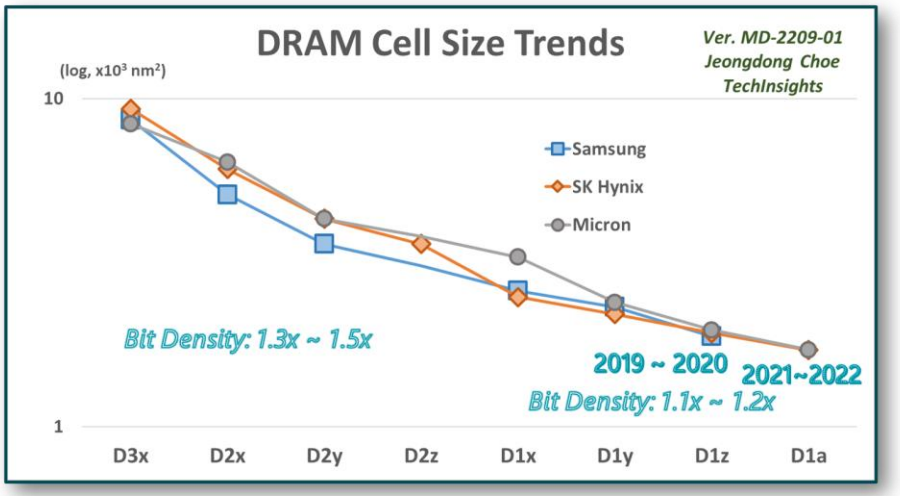

Figure 3. DRAM cell size trend: Samsung *vs.* SK Hynix *vs.* Micron

The high-k metal gate (HKMG) structure has firstly adopted for DRAM peripheral transistors for fully integrated and functioning DRAM from Samsung, and Micron just adopted it for GDDR6 peripheral transistors. The HKMG technology surely reduces gate leakage and improves DRAM performance. Gate leakages of HKMG are several orders lower than those of conventional Poly-Si/SiON structures for the same gate inversion oxide thickness (T_{inv}). The application of the HKMG technology will make DRAM gate dielectric scaling down by 10 Å or even less-10 Å regime. Samsung adopted SiGe/Si epi-layered channel on the LVP region to modulate the threshold voltage (V_{th}) by using band offset between SiGe and Si, while Micron keeps it with Si-channel without any band offset engineering on LV PMOS transistors. Further, La is not detected in LVN gate materials and interface layers (such as La_2O_3), which means Micron uses HfSiON/SiON gate oxide without any NMOS Vth modulation engineering. Samsung added a La-doped layer to

modulate the NMOS threshold voltage well. We know that La atoms diffuse and form dipoles at the interface and these dipoles decrease transistor V_{th}. Due to a potential abnormality in the TiN deposition profile ($LaCl_3$ formation), Micron might skip the use of La. Samsung successfully optimized the TiN deposition process for this, including O_2-less plasma PR strip process for HK dielectric (Table 1).

Table 1. A comparison of HKMG DRAM process and materials: Micron *vs.* Samsung

Device		Micron HKMG DRAM (Ex. D1z GDDR6)	Samsung HKMG DRAM (Ex. D1y DDR5)
Chips (Ex.)		MT61K512M32KPA-14:C (8 Gb)	K4RAH086VB-BCQK (16 Gb)
Die Markings		Y31G	K4RAH046VB
DRAM Cell Size		0.0020 μm²	0.0023 μm²
Bit Density		205.7 Mb/mm²	217.5 Mb/mm²
Peripheral Transistor Scheme		HKMG (Gate-First)	HKMG (Gate-First)
LVN	LVN TR Gox Materials & Thickness (POT), EOT	HfSiON (2.0 nm) / SiON (2.0 nm) (EOT ~ 1.5 nm)	La-doped HfSiON (2.0 nm) / SiON (2.0 nm) (EOT ~ 1.5 nm)
	LVN TR Gate Materials	W / WSiN / W/ TiN / Poly-Si / TiN	W/ TiN / Poly-Si / TiN 2
	LVN TR Gate Thickness (nm)	23 / 1.0 / 3.8 / 1.5 / 30 / 6.3	30 / 5.7 / 24 / 4.4
	LVN TR Gate Length (Observed)	51 nm	47 nm
LVP	LVP TR Gox Materials & Thickness (POT)	Al-doped HfSiON (2.0 nm) / SiON (2.0 nm) (EOT ~ 1.5 nm)	Al-doped HfSiON (2.0 nm) / SiON (2.0 nm) (EOT ~ 1.5 nm)
	LVP TR Gate Materials	W / WSiN / W/ TiN / Poly-Si / TiN	W/ TiN / Poly-Si / TiN 2 / TiN 1
	LVP TR Gate Thickness (nm)	23 / 1.0 / 3.8 / 1.5 / 30 / 6.3	30 / 5.7 / 24 / 4.4 / 3.0
	LVP TR Gate Length (Observed)	85 nm	78 nm
PMOS TR Channel		Si	SiGe-Epi

eFLASH is now replaced by eMRAM. A comparison of the Ambiq Apollo Blue MCU series is shown with Apollo1 through Apollo4 (Table 2). For reference, in comparison with Apollo3, peripheral and eMemory gate pitches have been decreased, 170 nm to 120 nm for the peripheral gate, and 230 nm to 110 nm for the eMRAM array for each. The supply current was scaled down from 35 uA/MHz for Apollo1 to 3 uA/MHz for Apollo4 eMRAM. All the Apollo Blue series have been supported by TSMC whatever it was eFLASH or eMRAM up to date.

Table 2. Ambiq Apollo Blue MCU series with TSMC eFLASH and eMRAM

Device	Ambiq Apollo Blue MCU	Ambiq Apollo2 Blue MCU	Ambiq Apollo3 Blue MCU	Ambiq Apollo4 Blue MCU
Product Example	Misfit Shine 2 Fitness	Huawei ERS-B29 Band 2 Pro	Huawei TER-B19 Band 3 Pro	Fitbit Luxe (Fitness Band)
Package Markings	APOLLO	AMAPH	AMA3B	AMA4B
Supply Current	35 μA/MHz (eFLASH)	10 μA/MHz (eFLASH)	6 μA/MHz (eFLASH)	3 μA/MHz (eMRAM)
Die Markings	ambiq apollo 1 2014	apollo 2 ambiq micro 2016	apollo 3 ambiq micro 2017	apollo 4 ambiq 2020
Die size (seal)	6.98 mm² (2.86 mm x 2.44 mm)	6.43 mm² (2.55 mm x 2.52 mm)	10.72 mm² (3.21 mm x 3.34 mm)	15.56 mm² (3.95 mm x 3.94 mm)
CMOS Process	90 nm	40 ULP	40 ULP	22 ULL
Foundry	TSMC	TSMC	TSMC	TSMC
Number of Metals	7 (6 Cu, 1 Al)	8 (7 Cu, 1 Al)	9 (8 Cu, 1 Al)	10 (9 Cu, 1 Al)
Logic Gate Pitch (min.)	740 nm	170 nm	170 nm	120 nm (HKMG)
Embedded Memory	eFLASH (2D NOR ESF3)	eFLASH (2D NOR ESF3)	eFLASH (2D NOR ESF3)	eMRAM
eMemory Gate Pitch	300 nm	250 nm	230 nm	110 nm
Bit Cell Size	0.087 μm²	0.068 μm²	0.068 μm²	0.046 μm²
eMemory Area Portion & Capacity	19.7 %, 4 Mb	14 %, 8 Mb	9 %, 8 Mb	10.4 %, 16 Mb

REFERENCES

[1] Jeongdong Choe, "Memory Technology 2021: Trends & Challenges," IEEE SISPAD 2021 Plenary Talk, 2021.
[2] Jeongdong Choe, "YMTC 128L Xtacking2.0 for SSD," Semiconductor Digest, December 2021
[3] Jeongdong Choe, "Micron 176L 3D NAND with 195T," Semiconductor Engineering, Aug. 19th, 2021
[4] Jeongdong Choe, "Micron D1α, The most advanced node yet on DRAM," Semiconductor Engineering, July 22, 2021
[5] Jeongdong Choe, "Samsung D1z DRAM with EUV Lithography," EETimes, Feb. 18, 2021