4 bits/cell 96 Layer Floating Gate 3D NAND with CMOS under Array Technology and SSDs

Pranav Kalavade

Intel Corporation, 2200 Mission College Blvd, Santa Clara, CA 95054 pranav.kalavade@intel.com

***Abstract*—This paper describes 4 bits/cell (QLC) 3D NAND based on 96 layer Floating Gate (FG) cell and CMOS under Array (CuA), achieving high areal density, performance, and reliability. More than 10million QLC FG 3D NAND SSDs have been shipped for both Client as well as Datacenter SSD applications, which is a significant milestone towards making 4 bits/cell NAND mainstream in the industry.**

I. Introduction

NAND Flash memory technology has undergone a major transition from 2D to 3D over the past decade. After many generations of (X,Y) feature size reduction to deliver orders of magnitude increase in die capacity, physical and electrical scaling constraints prompted the industry to make a disruptive change to 3D NAND so as to successfully continue the scaling trend [1-8] as shown in Fig. 1.

The intrinsic advantages from the larger physical size and Gate-All-Around structure of 3D NAND cells compared to their scaled 2D NAND counterparts have been well documented [1-10]. These include reduced number fluctuation effects resulting in tighter Vt distributions, low cell-to-cell interference, and good program disturb immunity due to large boosting ability. The superior cell properties in 3D NAND have enabled 3 bits/cell (TLC) to become mainstream.

The primary scaling knob for 3D NAND is to increase the number of layers, with more cells in each NAND string achieved through etching increasingly higher aspect-ratio memory holes and/or stacking multi-layer cell strings. However, these techniques lead to higher wafer manufacturing cost for 3D NAND as the number of layers increases through the technology generations. Increasing the number of bits/cell provides an orthogonal scaling path and offers a way to achieve cost reduction without building additional factory capacity. However, maintaining the NAND reliability at acceptable levels while increasing the number of bits/cell tends to be exponentially more challenging and requires robust cell characteristics. In fact, such transitions from N→N+1 bits/cell have happened only about once per decade through the history of mainstream Flash memory – 1 bit/cell (SLC) Flash invented: 1984, 2 bits/cell (MLC): ~1996, 3 bits/cell (TLC): ~2007. The intrinsic advantages of the 3D NAND Cell mentioned above naturally lends itself to such bits / cell scaling.

This paper highlights a 96 layer 4 bits/cell (QLC) FG- 3D NAND technology with CMOS Under Array (CuA) used for the manufacturing of the world's leading edge 1 Tb NAND Flash memory and incorporated into SSDs. The paper concludes with potential future directions of further extending scaling the FG 3D NAND technology to 5 bits / cell.

II. Technology

In transitioning from 2D to 3D, we had chosen a FG cell for its proven reliability. The first generation of 3D FG-NAND technology was introduced with 32 layers of active wordlines (WLs), plus additional tiers for dummy WLs, and source and drain select gates. [10]

The second generation doubled this to 64 layers, along with a number of key process innovations to engineer the cell stack, including inter-poly dielectric (IPD), FG, tunnel dielectric, and channel polysilicon, for higher program/erase window and better reliability throughout cycling and retention. Fig. 2 shows that >10V cell program/erase window with good program/erase slope is achieved. The large window is critical to accommodating twice the number of distinct Vt levels (from 8 to 16) needed to store 4 bits/cell compared to 3 bits/cell. [9]

The 64 layer process architecture was extended to 96 layers by increasing the number of active layers in each stack from 32 to 48. Fig. 3 shows the SEM cross sections of NAND array scaling methodology adopted for the successive generations of the FG 3D NAND utilizing stacking as well as increasing number of tiers.

NAND array engineering techniques are needed to overcome some of the reliability challenges posed by the migration from 3 bits/cell to 4 bits/cell. The effect of cell-to-cell interference on already placed and Vt-verified cells from programming neighboring cells is substantially reduced by using a multi-pass back and forth programming algorithm. In the first QLC product (64 layer), a 8-16 programming was used where the cells on a given wordline are programmed to intermediate Vt levels based on the data of the 3 LSBs while the MSB is used to program the final 16 level Vt distribution only after the neighboring wordline has seen most of its eventual Vt movement. In the subsequent 96 layer QLC, a 2-8-16 programming algorithm is used to reduce the cell-to-cell interference on the intermediate 8 level Vt distribution to improve the read window for wordlines in a partially programmed block. Fig. 4 shows the progression of the 16 level Vt placement. The larger Vt placement window for 4 bits/cell also leads to increased program, inhibit, and hot electron disturb. Several array voltage and timing waveform optimizations have

978-1-7281-6306-2/20/$31.00 ©2020 IEEE

been employed to manage channel electric fields and mitigate these sources of disturb.

The CuA technology was introduced in our first generation 3D NAND, enabling significant die size reduction. This first generation of CuA had the core circuits, such as sense amplifiers and decoders (wordline and bitline), placed underneath the NAND array. Starting with the second generation (64 layer) 3D-NAND technology and beyond, nearly all CMOS circuits are placed under the array, resulting in a very high array efficiency. Fig. 5 shows the result of the advancement in CuA that has been achieved due to layout optimization between the first and second generation 3 bits/cell FG 3D NAND. Additional improvements in CuA density enabled all additional circuits like extra page buffers, and negative WL voltage capability needed for supporting 4 bits/cell efficiently under the array, without compromising overall die efficiency.

In addition to die efficiency, the CuA technology enables read and write performance improvement. CuA enables a quad plane architecture where the high parallelism of 16KB page x 4 planes is used to enhance data throughput. In the 96 layer QLC FG 3D NAND, a new feature called Multi-plane read has been introduced, which allows two independent reads from the NAND die, significantly increasing the random read capability, again without sacrificing die efficiency.

Collectively, these innovations have resulted in the launch and commercialization of the first 1Tb 4 bits/cell NAND die in our second generation 64 layer FG 3D NAND in 2018 – a key milestone in NVM history, and was subsequently followed by the 96 layer FG 3D NAND. Fig. 6 shows measured 4 bits/cell Vt distributions. To our knowledge, this 96 layer based FG 3D NAND offers the highest areal density and performance among published 4 bits/cell NAND technologies (Table 1) [11,12,13].

III. 4 BITS/CELL SSDS

SSDs based on 4 bit/cell FG 3D NAND technologies addressing both Client as well as Datacenter applications have been developed. SSDs have stringent reliability and performance specifications, eg. meeting low annualized fail rate (AFR), total Terabytes Written (TBW), and data retention requirements >5 years while delivering high read/write bandwidth (MB/s, IOPS), and low latency with good Quality of Service (QoS). These specifications in turn translate into NAND component requirements for low defectivity, high cycling capability for all reliability mechanisms, while simultaneously meeting low read and program latency targets. The robust intrinsic cell characteristics of FG 3D NAND and the high bandwidth enabled by CuA, together with a well-engineered array enable 4 bit/cell FG 3D NAND to be an ideal candidate for SSD applications by delivering high storage capacity at lower cost.

A variety of SSD firmware and architecture innovations have been developed to further improve customer experience particularly with 4 bits/cell. In some cases, a novel variable-size SLC cache architecture is used to buffer host writes and enable faster performance associated with 1 bit/cell SLC NAND for typical usage cases while leveraging the higher density of the 4 bits/cell to also provide the benefits of higher storage capacity in the same SSD. Taking such concepts even further, hybrid SSDs combine Intel Optane technology with 4 bits/cell FG 3D NAND on a single M.2 2280 form factor. Intelligent driver software that caches frequently accessed data on the Optane media is utilized to deliver low latency and high performance on typical Client workloads, resulting in fast boot and application launch, in addition to smooth multitasking when such a hybrid SSD is used as an OS drive.

Over 10 million 4 bits/cell FG 3D NAND based SSDs have been shipped to date. This is a significant milestone towards making 4 bits/cell NAND mainstream in the industry. [14]

IV. POTENTIAL FUTURE DIRECTIONS

Toward scaling beyond 4 bits/cell, one can expect to leverage the benefits of the FG cell over the alternative Charge Trap Flash (CTF) cell that is also used in 3D NAND. Fig. 7 shows a comparison of the data retention capability of 4 bits/cell 3D NAND made using the FG cell and CTF cell respectively. As reported in the literature also [15,16], the FG cell clearly shows superior data retention properties, attributed to one of the key differences between FG and CTF: the fact that the charge storage layer is continuous in CTF while is it isolated in FG. Fig. 8 shows 5 bits/cell Vt distributions demonstrated on FG 3D NAND. This is one of the attractive directions that could be pursued towards displacing hard disk drives with solid-state storage.

V. CONCLUSION

4 bits/cell technology using Floating Gate 3D NAND technology and CMOS Under Array architecture has been developed offering high endurance and reliability while delivering high areal density and performance. 4 bits/cell FG 3D NAND SSDs have been shipped for both Client and Datacenter applications. The FG 3D NAND cell shows promise to enable even further scaling to 5 bits/cells.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of the large teams from Intel as well as Micron towards obtaining the results shown in this paper.

REFERENCES

[1] J-W. Im, et. al, *ISSCC Tech Digest*, pp. 130 - 132, 2015.
[2] R. Yamashita, et. al, *ISSCC Tech Digest*, pp. 130 - 132, 2017.
[3] H. Tanaka et al., *VLSI Tech Symp. Digest*, pp. 14-15, 2007.
[4] J. Jang, et al., *VLSI Tech Symp. Digest*, pp. 192-193, 2009.
[5] R. Katsumata, et al., *VLSI Tech Symp. Digest*, pp. 136-137, 2009.
[6] H-T. Liu, *VLSI Tech Symp. Digest*, pp. 131-132, 2010.
[7] J. Lee, et. al, *IEDM Tech Digest*, pp. 284 - 287, 2016.
[8] S. Lee, et. al, *ISSCC Tech Digest*, pp. 340 - 342, 2018.
[9] K. Parat, et. al, *IEDM Tech Digest*, pp. 27 - 20, 2018.
[10] K. Parat, et. al, *IEDM Tech Digest*, pp. 48 - 51, 2015.
[11] N. Shibata, et. al, *ISSCC Tech Digest*, pp. 210 - 210, 2019.
[12] D-H. Kim, et. al, *ISSCC Tech Digest*, pp. 218 - 219, 2020
[13] H. Huh, et. al, *ISSCC Tech Digest*, pp. 220 - 221, 2020
[14] Press release: https://newsroom.intel.com/articles/intel-builds-10-million-qlc-3d-nand-solid-state-drives/#gs.3t8sog
[15] B. Choi, et. al, *VLSI Tech Digest*, pp. 78 - 79, 2016
[16] C-P. Chen, et. al, *IEDM Tech Digest*, pp. 118 - 121, 2010

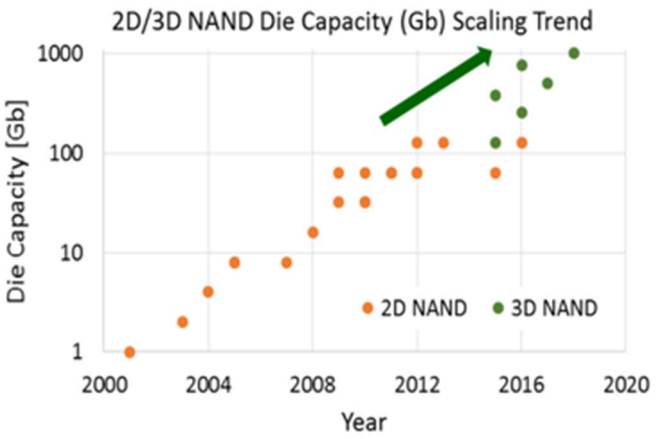

Fig. 1. Transition to 3D NAND has extended the die capacity scaling to 1Tb

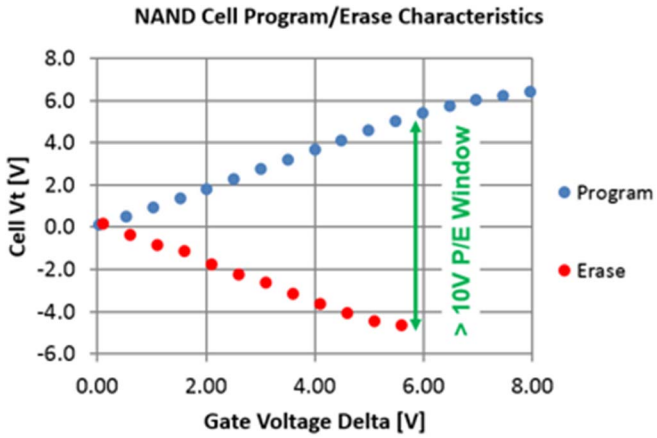

Fig. 2. NAND Cell Program / Erase characteristics

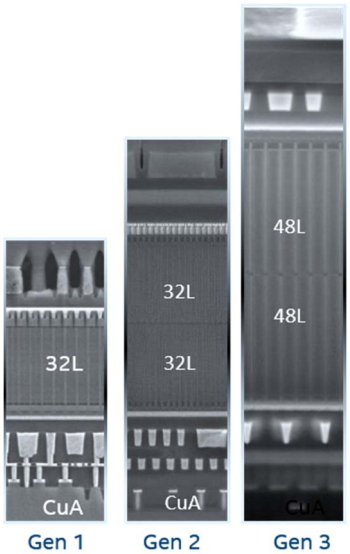

Fig. 3. SEM cross sections of FG 3D NAND with 32, 64, and 96 active layers and CMOS Under Array

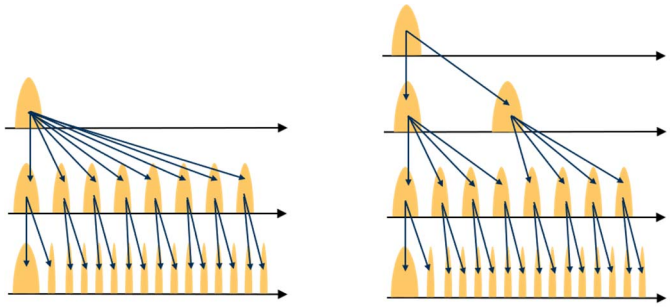

Fig. 4. 16 Level placement using 8-16 (64 layer) and 2-8-16 (96 layer) programming algorithm

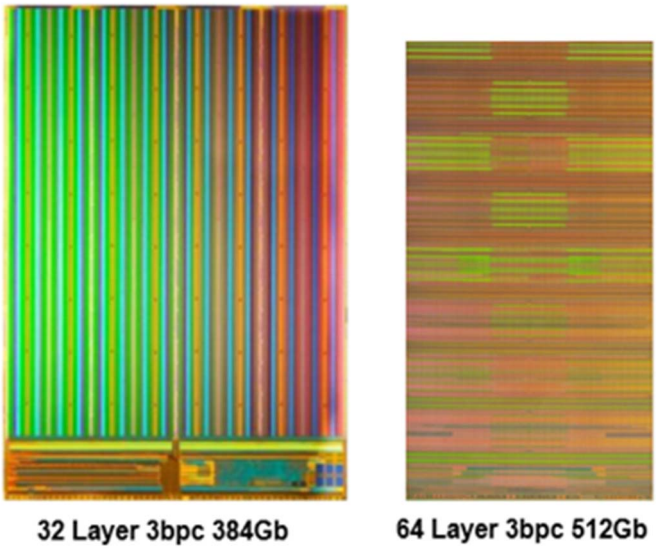

Fig. 5. Layout improvement between 1st and 2nd Gen to fit majority of circuits under the NAND Array.

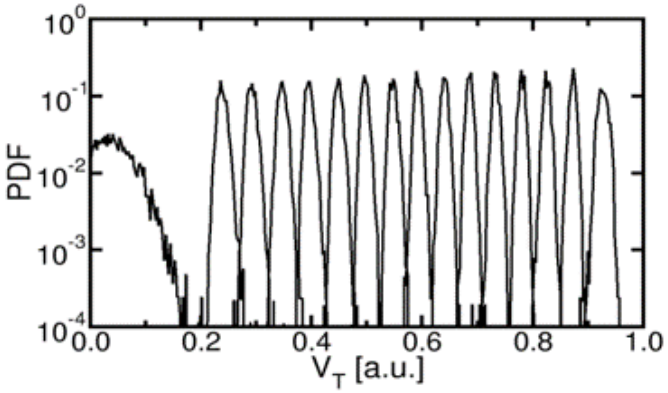

Fig. 6. Vt distributions of cells in the 4b/c die

	Intel [This paper]	Ref [11]	Ref [12]	Ref [13]
Active Tiers (#)	96	96	92	96
Bit Density (Gb/mm²)	8.9	8.5	7.5	8.4
Seq Write Perf (MB/s/die)	30	9.3	15.6	30

Table 1. Comparison of Areal Density and Write Performance for various 4 bit/cell die's

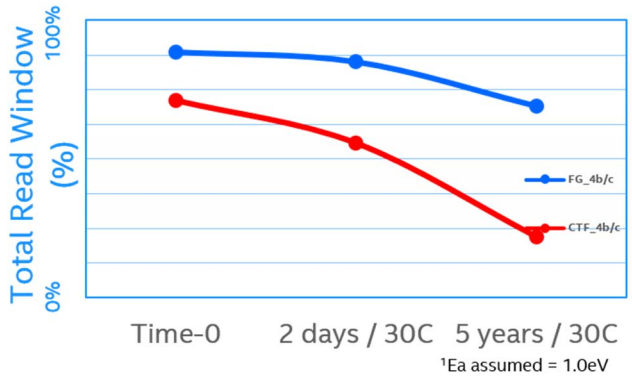

Fig. 7. Floating Gate Cell advantage for data retention

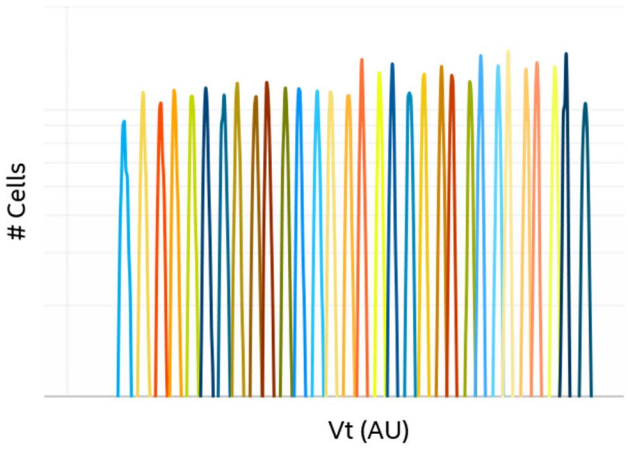

Fig. 8. Intel FG 3D NAND shows promise for 5 bits/cell