3D NAND Technology – Implications to Enterprise Storage Applications

Jung H. Yoon

Memory Technology
IBM Systems Supply Chain
Outline

- Memory Technology Scaling - Driving Forces
  - Density trends & outlook
  - Bit cost factors
- 3D NAND Design & Architecture key factors
- 3D NAND Process, Reliability & Quality
- 3D NAND Performance, Power
- 3D NAND Technology – Implications to Enterprise Storage
- Summary
- 2D NAND continues to drive lithographic minimum feature scaling – 1y/1z nm in volume production
- Floating gate scaling to 1z nm (~15nm) > industry wide transition to 3D NAND cell in 1H16 enabling path for ‘Effective Sub 10nm’ scaling
- Timing of 3D NAND implementation to Enterprise Storage will depend on 3D MLC/TLC Flash technology yield, reliability maturity, bit cost reduction, combined with Controller/Flash management enablement in 2016-2017 timeframe
- TLC currently accounts for ~40% of industry output, anticipate to exceed 65% by 2018
  - 3D TLC will drive continued bit cost reduction & opportunities for wider flash adaptation including enterprise storage
DRAM & Flash Scaling – Density Trends

DRAM scaling at sub 20nm node for 12/16Gbit – bit cost reduction vs increased technology complexity & fab investment requirements

Flash scaling continues via 2D > 3D NAND transition – enabling 256Gb & higher density MLC/TLC Flash
Memory Bit Cost Curve vs Scaling

- Steep increase in Fab Capital Expenditure @ sub 20 nm – driven by immersion ArF tooling & low K1 lithography, multiple patterning overhead
- Flash Bit cost reduction will continue at steep rate via 3D NAND scaling in 2016-2020 – 3D NAND yield, quality and reliability is key
- DRAM $ per GB take down slope significantly flattened at 1x/1y/1z nm in 2016-2019
3D NAND Bit Cost

Minimal cell size ~ 4Xnm 6F²

1. 3D NAND bit cost (simple model) \( \propto \frac{1}{N} \)

2. Additional Staircase Contact adds process cost
3. Added Area for Decoder – WL decoder area needed to reduce RC delay
4. Added Process cost for peripheral circuit integration
5. Bit cost reduction @ 64 - 100 layers will required vertical channel etch profile, innovation in decoder/peripheral design/layout, F & z-directional wordline-to-wordline pitch reduction
6. Fab wafer ramp-up, yield & quality maturity key in 3D NAND bit cost. 3D NAND process specific Fab CaPex needed for initial production, Fab thruput key factor for bit cost

Source: PI Du ICSICT 2014

Source: A. Walker, IEEE2013
3D NAND Design & Architecture Key Factors

1. 3D NAND Cell array architecture: Page, block, plane size & structure
2. Program & Read algorithms
3. X/Y and Z directional cell-to-cell interference
4. ECC requirements & error characteristics
3D NAND P/E Cycling

3D Charge Trap – Key Characteristics

1. Reduction in Cell-to-cell interference due to lithographic cell spacing relaxation (~15 nm > 4x nm)
2. Charge Trap Thin Tunnel Oxide – less charge trap build up caused by PE Cycling => tight Vt distribution
3. Enables faster programming speed due with 1 pass programming algorithm
3D Charge Trap – Data Retention

- Fast initial charge loss due to shallow trapped electrons
- Data Retention fails - due to charge spreading across channel (Z-direction) and charge loss thru thin tunnel oxide (X-direction)
- Understanding of High Temp & Low Temperature data retention mechanism critical – degraded data retention characteristics at > 125C anticipated
- 3D NAND Floating gate cell expected to have an advantage in data retention but with endurance characteristics tradeoffs
1. High Aspect Ratio Channel Etch Profile (Distortion free, near vertical θ angle) & ONO film thickness control in z-direction
2. Issues at bottom gate has ‘ripple effect’ all the way up the channel
3. Staircase contacts from the Word-line drivers for each layer creates cell structure unique to 3D NAND – High A/R contact etch, alignment accuracy, contact resistance uniformity
4. Bit line contact @ top of channel – interface properties & alignment accuracy
5. 3D NAND specific Defect control, metrology & methodologies needed
6. Wafer yield – Drive wafer to wafer, across wafer, die-to-die, intra-die variability reduction => critical for 3D NAND Bit cost and Enterprise Quality & Reliability
3D NAND Cell/Process Architecture Challenges

- Wordline RC delay – more Wordline decoder area needed to relax large RC Loading
- Wordline capacitance increase (due to 3D NAND architecture) - Icc current increases
- Low cell current & variability due to poly-Si channel – innovations for high mobility channel needed for continued scaling of 3D NAND layer count, Poly Si microstructure engineering critical
- Random Telegraph Noise (RTN) – charge trap, carrier mobility fluctuation

Techinsights Samsung V-NAND 2014

M. Toledano et al., IEDM 2013
3D NAND Programming Scheme

- 3D NAND allows fast programming speed due to 1 pass Programming algorithm – possible with reduced Cell-to-Cell Interference
- Dual/triple pages can be programmed simultaneously with less programming steps compared to 2D NAND
- Reduction in tPROG > Reduction in latency > Higher performance
- BER reduction via elimination of Partial Page Programming/upper page read error scheme
3D NAND Performance & Power Consumption

Program time & Energy Consumption

- 3D NAND allows 1 pass Programming Algorithm => tPROG reduction
- Potential for Energy Consumption reduction – via reduced tPROG
  - Normalized Energy Consumption = Vcc x ICC2 x tPROG / Page size
  - Possible ICC current increases due to 3D NAND architecture specifics
- Potential for Reduction in Flash Operation power and Sequential Write Power Efficiency => positive for OpEx related power/cooling costs

Source: J.W Im et.al. ISSCC 2015

Flash Memory Summit 2015
3D NAND Technology – Implications to Enterprise Storage

- **Density**
  - ✓ 256Gb+ MLC/TLC Flash density with same footprint package

- **Reliability**
  - ✓ Driven by lower Cell-to-Cell Interference & Tighter Vt distribution with P/E cycling – 3D Charge Trap vs 3D Floating gate needs to be further evaluated/understood.
  - ✓ Flash characterization critical
  - ✓ Data Retention for 3D Charge Trap anticipated to worsen – need thorough evaluation & understanding of mechanism over wide operational & storage temperature range

- **Performance**
  - ✓ Faster tPROG due to 1 pass programming algorithm – due to reduced Cell-to-Cell interference
  - ✓ Dual/Triple pages simultaneous programming with less programming steps

- **Power Efficiency**
  - ✓ Potential Power reduction driven by tPROG reduction
  - ✓ ICC current tradeoffs need to be understood for overall Power efficiency gains
4 ‘V’s of Data – Key Considerations for Big Data

**VALUE**
- Rapidly declining Flash Bit Cost (MLC/TLC) via 3D NAND scaling
- Optimize economic value of data – fast analysis of more data in different formats
- OpEx benefits – Low power/cooling costs
- Less data center physical space due to higher density flash

**VOLUME**
- Volume of data captured, stored, shared, and analyzed continue to increase
- Raw capacity growth
- Greater storage density in the data center

**VARIETY**
- Unstructured data – files, email, social media consists >80% of all data
- Accommodate file & object storage represented by unstructured data – best suited for Flash storage

**VELOCITY**
- Performance – data must be analyzed faster > smarter decisions & maximize business value
- Higher IOPS, lower latency needed

3D NAND critical to Enterprise Storage - continue to drive Flash bit cost reduction, increases in Capacity, Reliability, Performance and Reduction in power consumption
Summary

- 3D NAND introduction and scaling – anticipate continued MLC/TLC bit cost reduction while driving Density growth over the next 5+ years. Bit cost scaling will be determined by Si process, device and design innovations focused on channel profile, vertical transistor, staircase process controls, decoder & peripheral circuit complexity, and process uniformity.

- Timing of 3D NAND implementation to Enterprise Storage will depend on 3D MLC/TLC Flash technology yield, reliability maturity, bit cost reduction, combined with Controller/Flash management enablement in 2016-2017 timeframe.

- TLC current accounts for ~40% of industry output, anticipate to exceed 65% by 2018.
  - 3D TLC will drive continued bit cost reduction & opportunities for wider flash adaptation including enterprise storage.

- 3D NAND cell architecture
  - Relaxation of lithographic pitch reduces Cell-to-Cell interference & tighter Vt distribution => Endurance & Performance gain opportunities anticipated.
  - Data retention expected to be weaker 3D Charge Trap – need further understanding of data retention at operating & storage temperatures.

- 3D NAND Process Technology unique challenges – fab quality process controls critical.

- 1 pass programming allows shorter tPROG => opportunity for performance gain & power consumption reduction.

- 3D NAND Technology critical for Enterprise Flash Storage - driving Flash bit cost reduction, increases in Capacity, Reliability, and opportunity for Performance gain & OpEx benefits via Power consumption reduction.