IBM

3D-NAND scaling & 3D-SCM – Implications to Enterprise Storage

Jung H. Yoon, Ranjana Godse
IBM Supply Chain Engineering

Gary Tressler
IBM Flash SSD Development

Hillery Hunter
IBM Research

© 2017 IBM Corporation

The Current Storage Model is being Disrupted by the Explosion of Data and Need for Speed

Data Explosion

- 2.5 Billion Gigabytes of data per day
- 90% of data created in last two years

Data Economics

- 40% more data per year for storage administrator
- Flat overall IT budget

Data Innovation

- 30% lower TCO with Flash
- 50% lower storage management cost with software defined storage

Cognitive Computing will change the business & needs of Data Storage

- The digital world generating oceans of data which must be efficiently stored, managed, and protected as well as used by the right applications at the right time
- Storage Model disruption driven by Si technology innovations: 3D-NAND and 3D-SCM technologies

© 2017 IBM Corporation

Enterprise Storage Market Trends and Transitions

1. Enterprise Flash-SSD Exabyte Growth

- 3D NAND enabling 512Gb /1Tb densities in 2017-2018
- 3D-TLC density and reliability allowing strong penetration to enterprise storage, read intensive applications
- Stacked packaging – 8DP/16DP/32DP, TSV

2. 3D NAND Scaling

- 3D NAND scaling allowing path for density growth, bit cost reduction to 2022+
- Interest in 3D-QLC for further bit cost reduction, focused on Cold storage, TLC/QLC Tiering

3. NVMe Adoption in Server and Storage systems

- NVMe based storage systems expected to enable significantly lower latencies - speed data to and from storage solutions and systems
- PCIe Gen4 provides significant processor ability to absorb data, acceleration expected in 2018

4. Storage Class Memory enablement

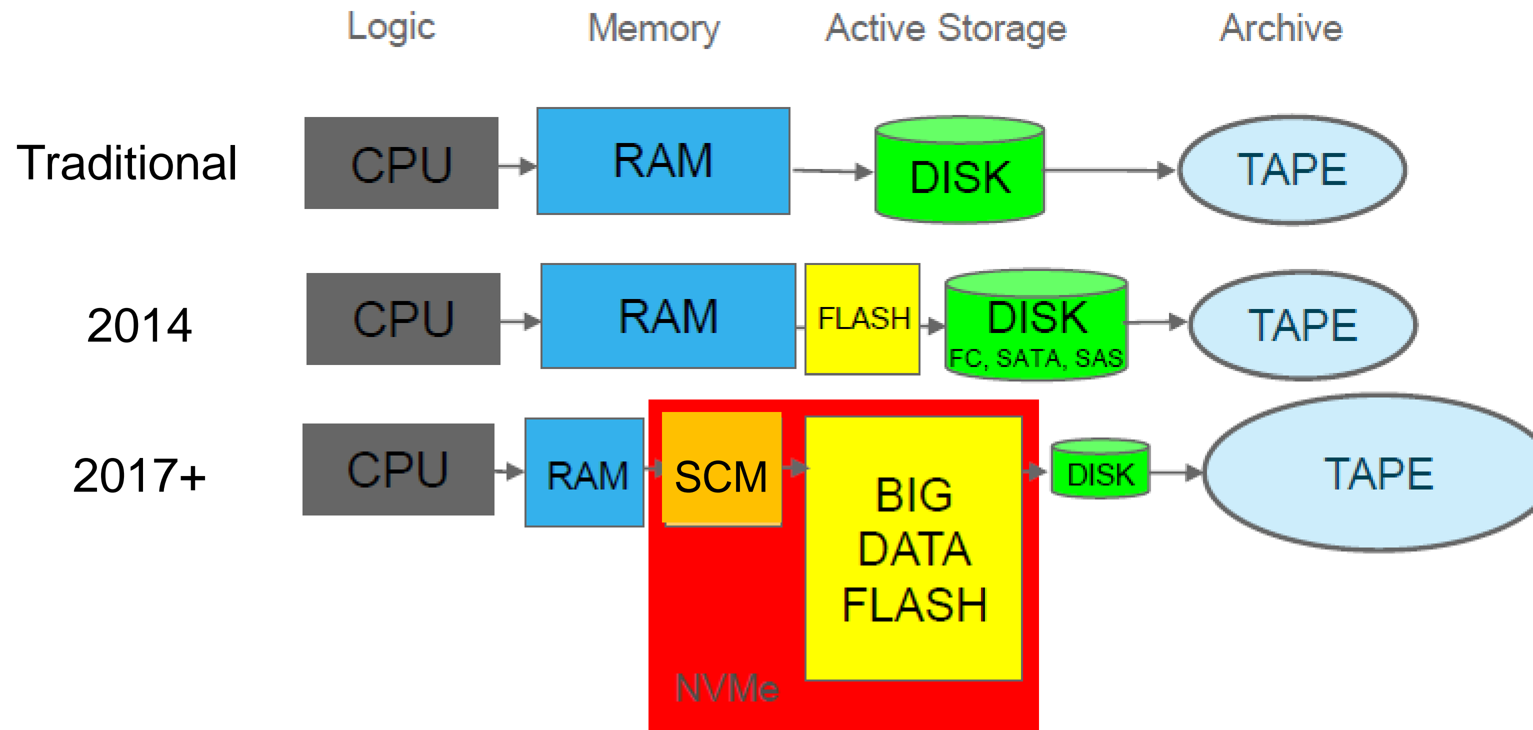
- Multiple deployment options under exploration
- PCIe NVMe block attach for initial time to market
- PCIe memory mapped, memory bus attach for highest performance, SCM memory assessable via Open Capi high bandwidth, low latency I/F

5. Development of "Flash" Focused form factor

- New SSD form factors optimized for flash deployment
- 3D-NAND die, package form factors, Gb/mm^3 critical

© 2017 IBM Corporation

Evolution of Storage and Memory Hierarchy

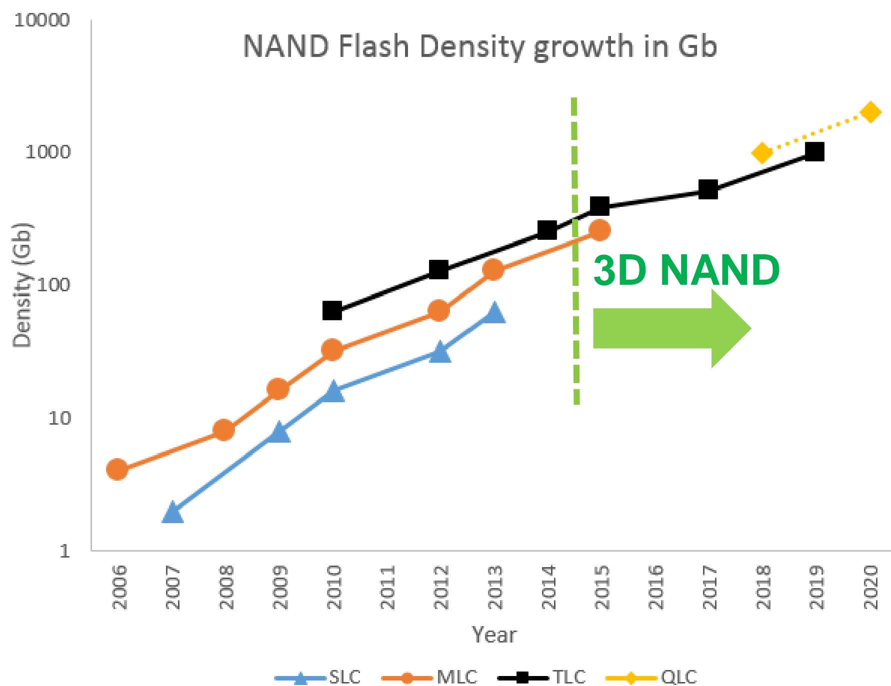


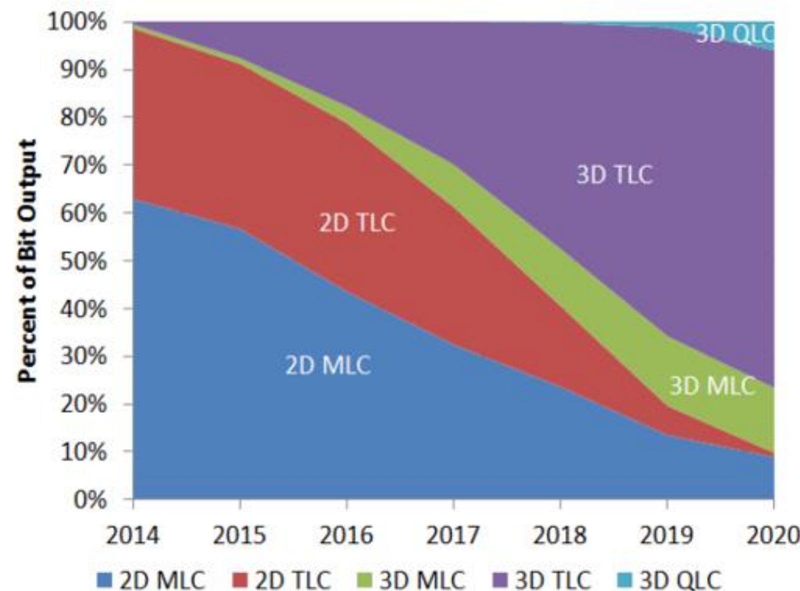
A new category of 'Big Data Flash'

- Many workloads do not require the write performance and endurance of Enterprise Flash
- Priorities are centered at high density, low cost, and good read performance – enablement driven by 3D-TLC & QLC scaling

© 2017 IBM Corporation

Flash Density Growth & Technology Transition





Source- Forward Insights

- Flash scaling continues via 3D NAND layer count increase – enabling 512Gb~1Tb & higher density TLC /QLC Flash in 2017-2018
- Cost per bit reduction will continue via 3D NAND scaling 2017-2022+
  - 3D-NAND scaling provides clear path for flash density growth – 512Gb 3D-TLC in 2H17 (64L), 1Tb QLC in 2018
  - 3D NAND scaling expected to continue on a 18 months cadence thru 2020 – 48L >64L > 96L >120L+
  - 3D-NAND (TLC/QLC) $/GB cost take down delivering Flash TCO advantages

© 2017 IBM Corporation

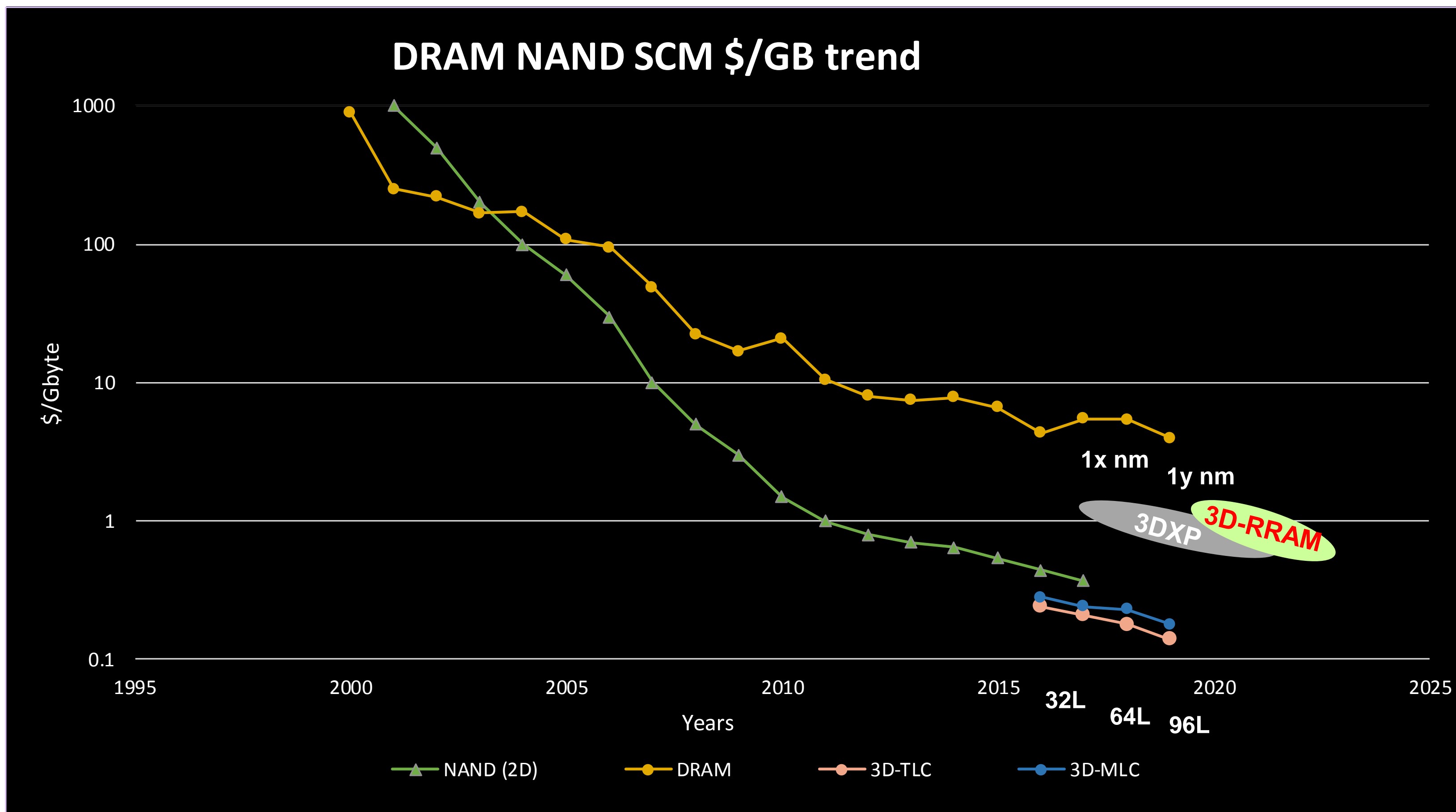
3D NAND Production Transition

Industry wide transition 3D NAND enabling 512Gb+ densities in 2H17 with significant fab CAPEX investments

- 3D NAND's improved cell reliability enables TLC usage for enterprise storage applications, while driving bit cost reduction – strong penetration of 3D TLC in Enterprise SSD/Storage
- Significant fab investments ongoing for new flash Fabs in 2017-18
- TLC current accounts for ~50% industry output, will exceed >80% in 2019
- QLC introduction in 2018 at 1Tb+ densities, ramp up in 2019
  - Focus on Archive & Cold Storage Applications – targeting Near Line HDD market
  - QLC/TLC Tiering applications

© 2017 IBM Corporation

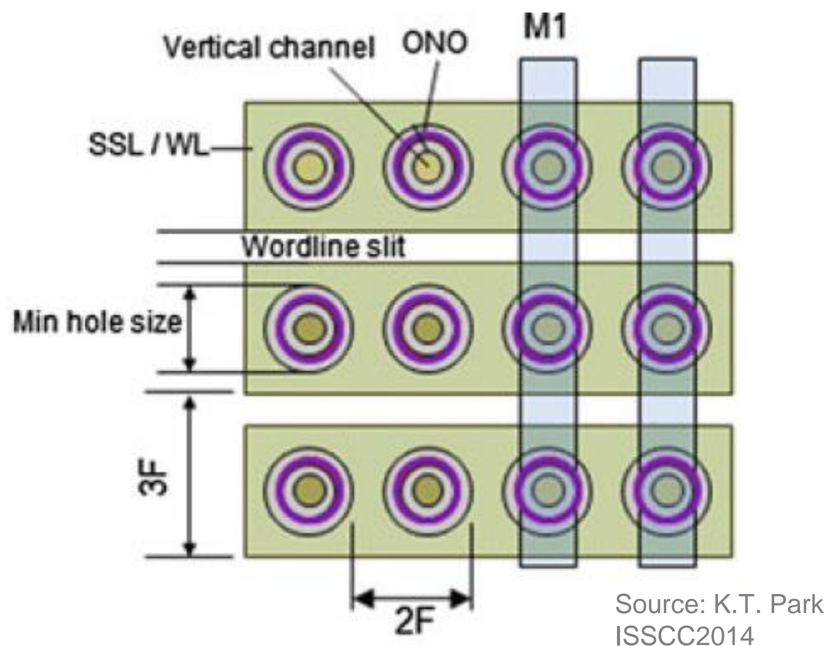
3D NAND vs 3D-NVM Cost Projections



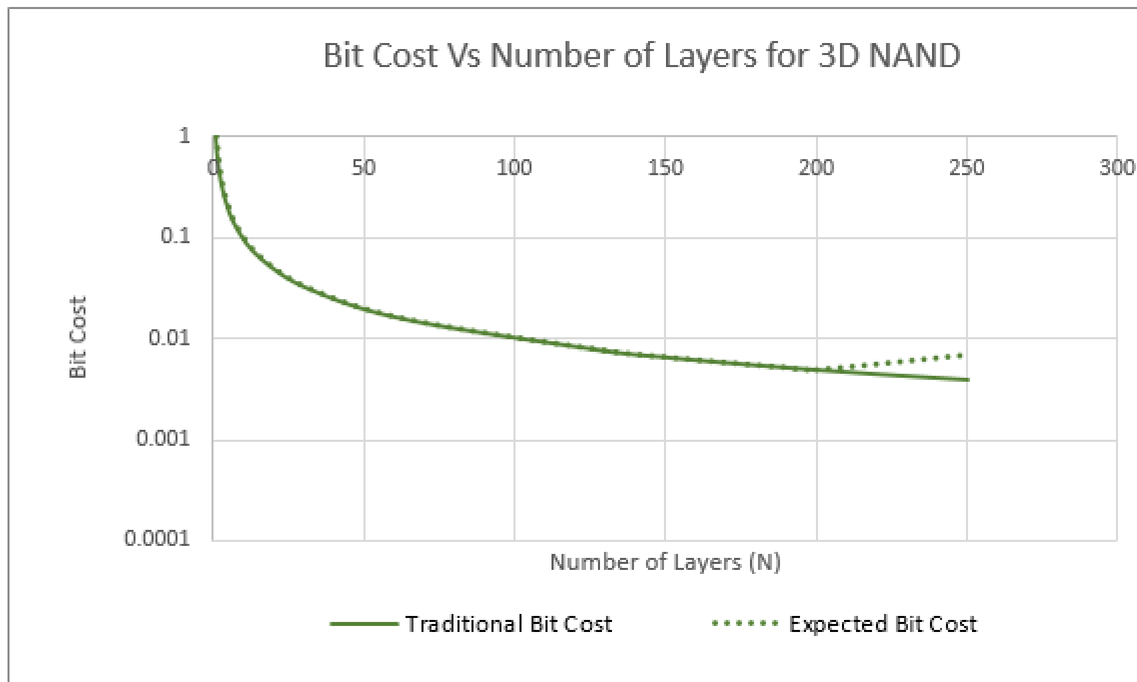
- Exponential Flash growth anticipated thru 2020, driven by 3D-NAND scaling, scaling cadence @ 18 months/generation
- Wide usage of 3D TLC in Cloud Datacenter & Enterprise Server Applications starting in 1H17.
- 3DXP and 3D-ReRAM $/GB positioned between DRAM & NAND, slope will depend on scaling approaches & technologies

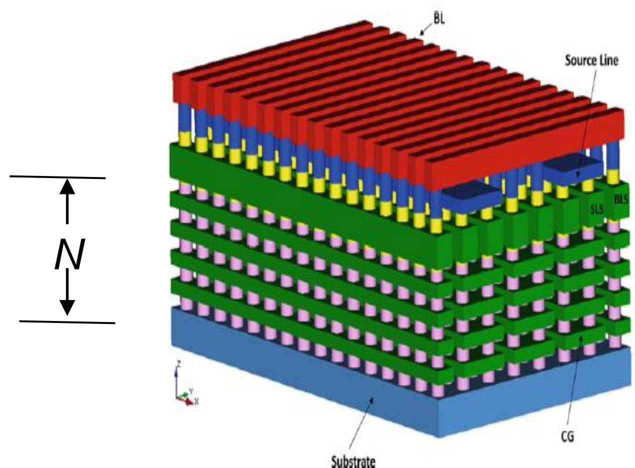
© 2015 IBM Corporation

3D NAND Layer Count Scaling vs Bit Cost



Source: K.T. Park ISSCC2014



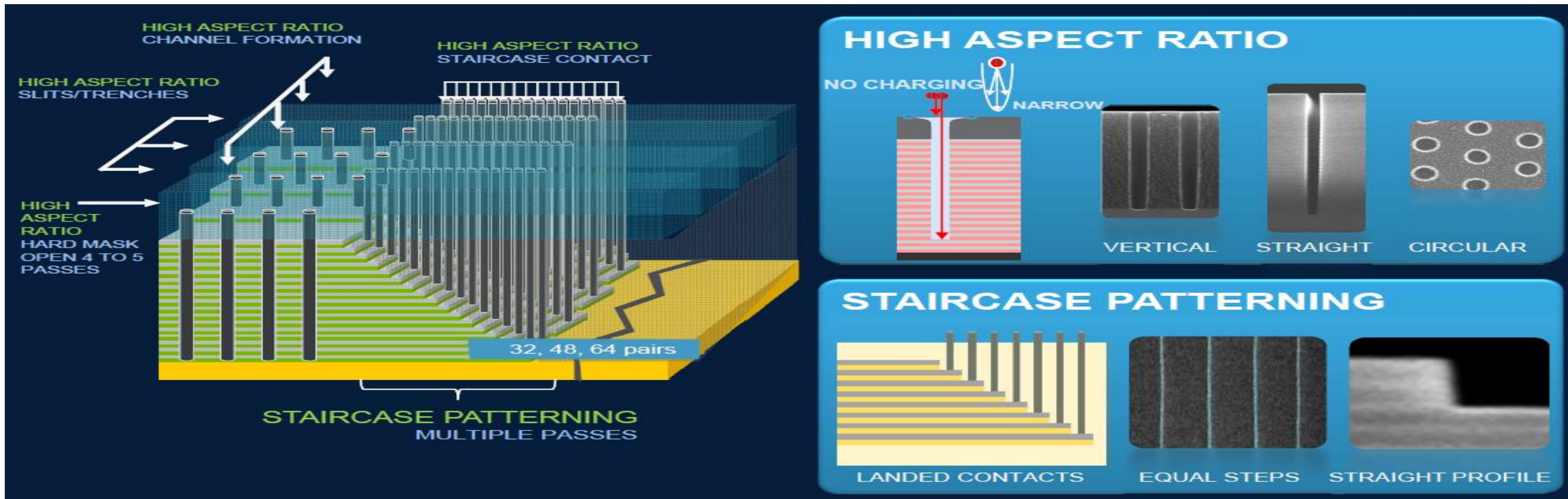


1. 3D NAND bit cost (simple model) $\propto \frac{1}{N}$
2. Additional Staircase Contact adds process cost
3. Added Area for Decoder – WL decoder area needed to reduce RC delay
4. Bit cost reduction > 120 layers will require vertical channel etch, innovation in decoder/peripheral design /layout, F & z-directional wordline-to-wordline pitch reduction
5. Fab wafer ramp-up, edge die yield & quality maturity key in 3D NAND bit cost.

© 2017 IBM Corporation

3D NAND Scaling Trends, Outlook & Challenges

	2016	2017	2018	2019	2020
3D NAND Generation	32/48L	64L	64/96L	96L	96/128L
Max density (TLC/QLC)	256Gb+	512Gb	1Tb	1Tb+	1-2Tb+



Source: Applied Materials, IEDM, VLSI

**** Process challenges anticipated with** <u>**> 120 layer 3D NAND scaling**</u> **– process innovations required in High AR multiple pass Channel Etch , Stair case contact process, defect control, wafer edge variability & quality, wafer yield***

© 2017 IBM Corporation

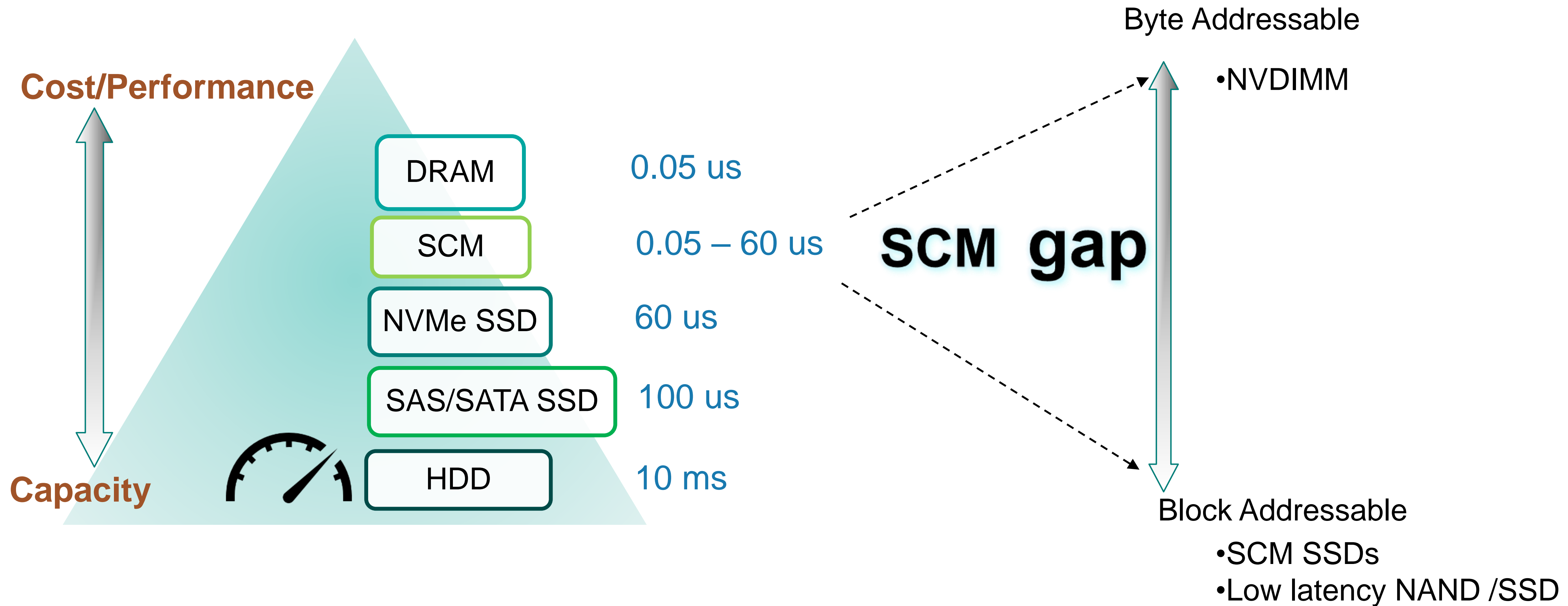
3D NAND Technology – Implications to Enterprise Storage

- **Density**
  - ✓ 512Gb 3D-TLC, 1Tb 3D-QLC Flash density in 2017-2018
- **Reliability**
  - ✓ Driven by lower Cell-to-Cell Interference & Tighter V_t distribution with P/E cycling
  - ✓ Sustained reliability (Endurance, Data Retention, Read disturb etc.) characteristics with 3D-NAND scaling (48L > 64L> 96L…)
  - ✓ QLC density & bit cost advantages vs reliability tradeoff
- **Performance**
  - ✓ Faster t_{PROG} due to 1 pass programming algorithm – due to reduced Cell-to-Cell interference, Dual/Triple pages simultaneous programming with less programming steps
  - ✓ Sustained/improved performance parametrics (t_{Read}, t_{PROG}, t_{BERASE} etc.) with 3D NAND scaling
  - ✓ QLC performance vs density/cost tradeoffs need to be understood for TLC replacement consideration
- **Power Efficiency**
  - ✓ Potential Power reduction driven by t_{PROG} reduction
  - ✓ Additional Power reduction by additional plan programming (2 vs 4 plane)
  - ✓ High density 3D-TLC enables flash based system’s lower TCO

© 2017 IBM Corporation

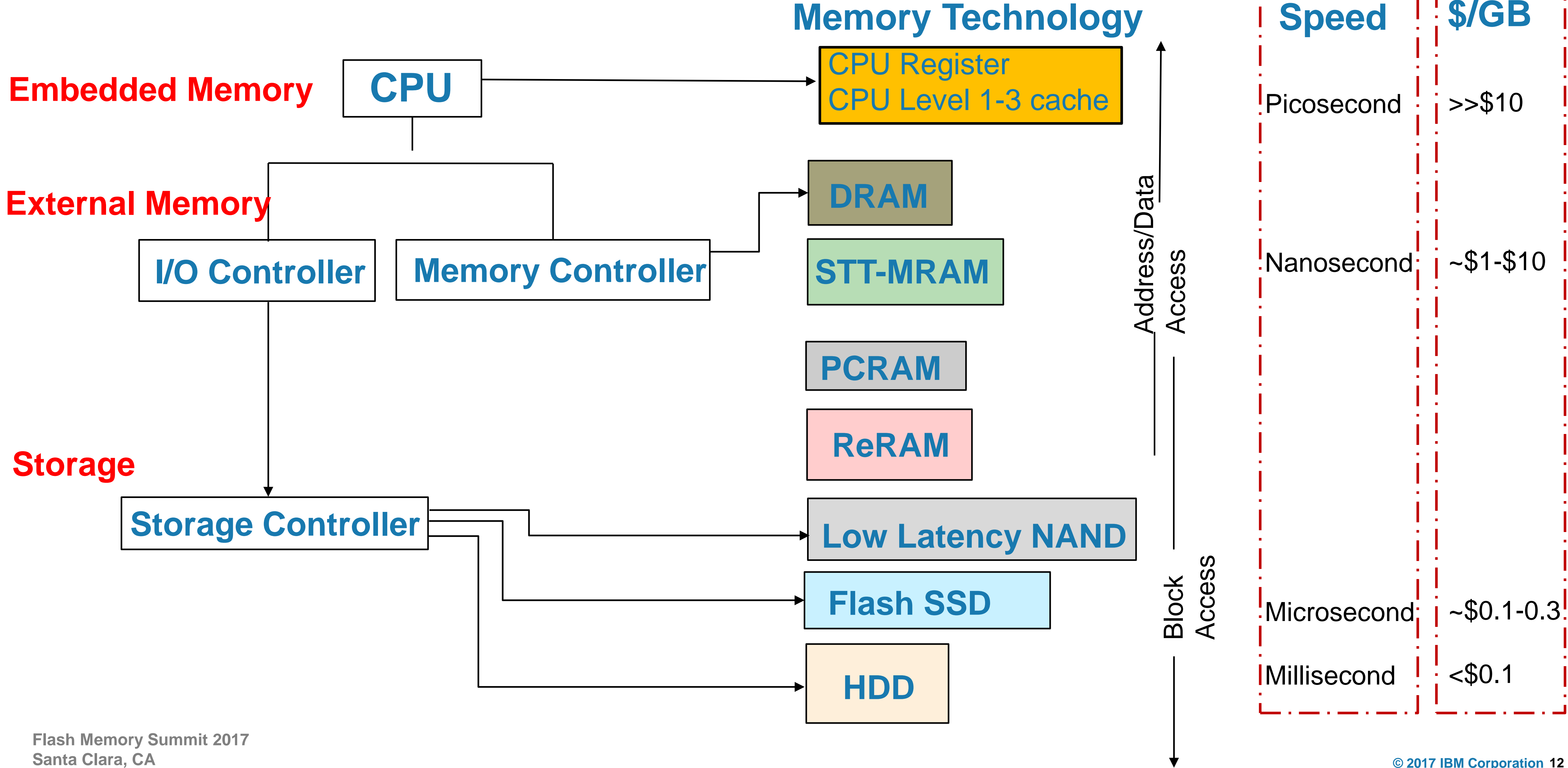
Storage Class Memory Technology Implication

Memory and Storage Hierarchy



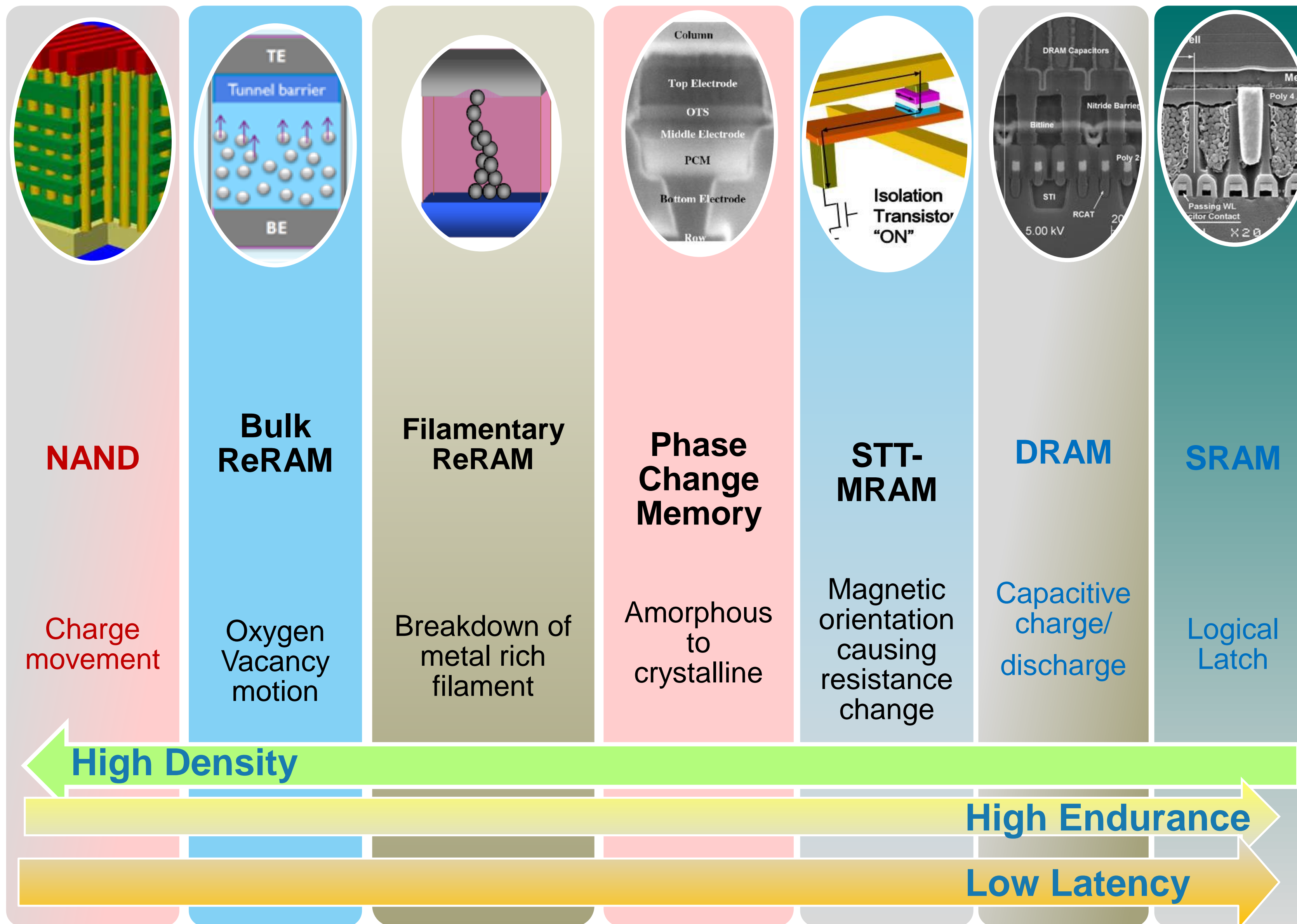
© 2017 IBM Corporation

System Architecture - Memory Tiers



© 2017 IBM Corporation

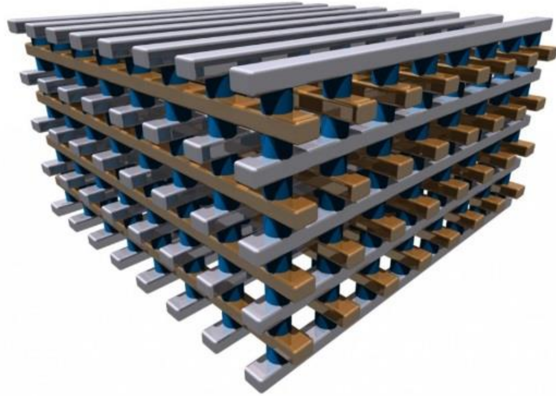
Storage Class Memory – Key Comparison



Memory Type	NAND	Bulk ReRAM	Filamentary ReRAM	Phase Change Memory	STT-MRAM	DRAM	SRAM
Switching Mechanism	Charge movement	Oxygen Vacancy motion	Breakdown of metal rich filament	Amorphous to crystalline	Magnetic orientation causing resistance change	Capacitive charge/ discharge	Logical Latch

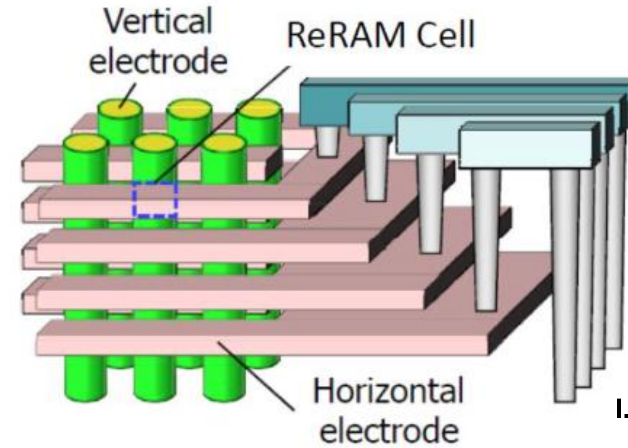
© 2017 IBM Corporation

3D- ReRAM Cell Comparisons



Horizontal ReRAM

- Number of critical masks increase with added number of layers
- Requires Advance Lithography
- ReRAM cell principle is same as Vertical ReRAM



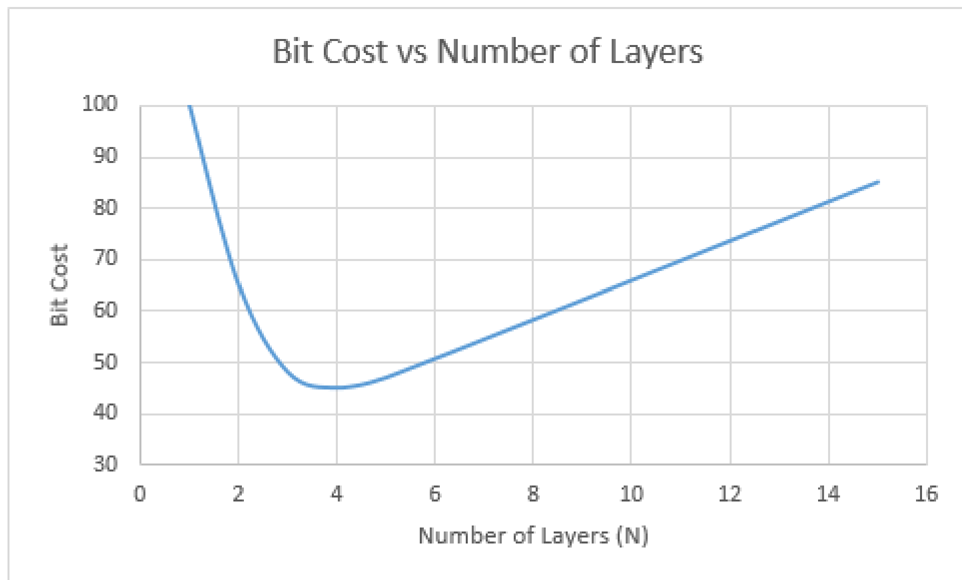
I. Baek, et. al, IEDM2014

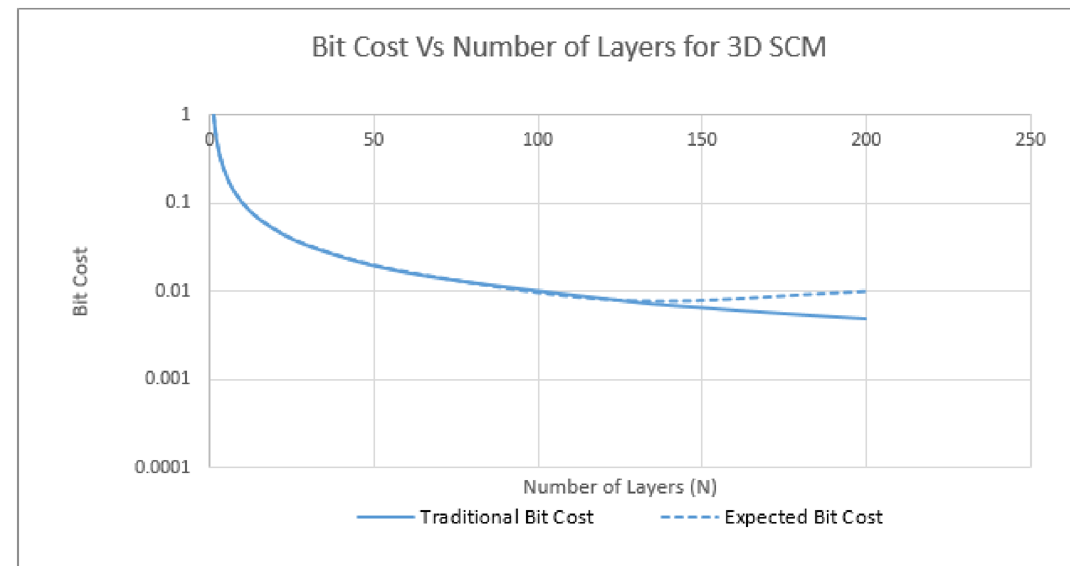
Vertical ReRAM

- Number of critical masks do not increase with added number of layers
- 3D NAND like process
- Challenging to integrate the selector

© 2017 IBM Corporation

3D- SCM Scaling – Bit Cost Key Factors



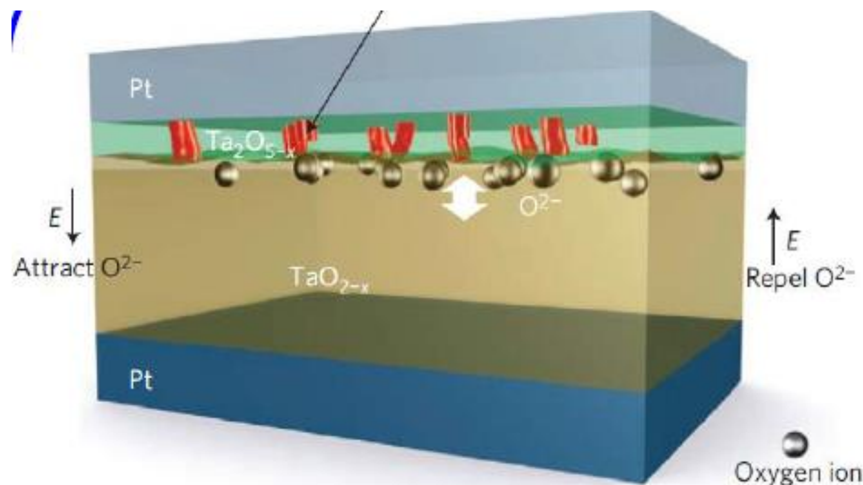


- Cross Point 3D-ReRAM scalability strongly lithography dependent – due to complexity/cost associated with high layer count 3D cell
- Vertical type 3D-ReRAM has a more complex cell structure – switching and cell variability challenging. Long term scalability and cost reduction viewed positive

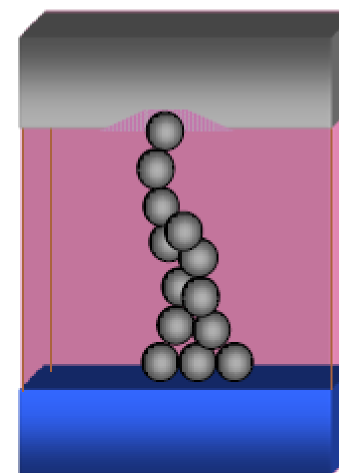
© 2017 IBM Corporation

3D-ReRAM Reliability and Manufacturing challenges

- Endurance – dependent on consistency of filament formation
- Data retention – highly temperature dependent, discontinuity in conductive filament due to oxidation and diffusion effects
- Sneak path current – leakage current path reducing resistance of 3D-ReRAM memory cell
- Cell variability – significant challenge due to physics that depends stochastic nature of formation and dissolution of conductive filament



Resistive Switching MeOx ReRAM



Conductive Bridge ReRAM

© 2017 IBM Corporation

Summary

- Current Storage Model is being disrupted by the explosion of data and need for speed
- Cognitive Computing changing the business & needs of data storage - Storage Model disruption driven by Si technology innovations of 3D-NAND and 3D-SCM technologies & enablement
- Enterprise Storage driving need for Big Data Flash – enabled by 3D-TLC/QLC’s density, low cost, read performance
- Clear roadmap for 3D-NAND scaling into 2022+ timeframe, with layer counts exceeding > 250 Layers, Process technology & design innovations will need to continue for continued $/GB reduction with scaling
- 3D-ReRAM focus as memory tier between 3D NAND and 3DXP – technology enablement and maturity will need to continue for enabling densities into ranges that are interesting to enterprise storage

© 2017 IBM Corporation