Is 2017 the Year of Emerging Memories?

Dr. Seshubabu Desu
Chief Technology Officer
4DS Memory Limited
Fremont, CA
sbdesu@gmail.com

FMS-August 2017

DRAM Scaling Limitations

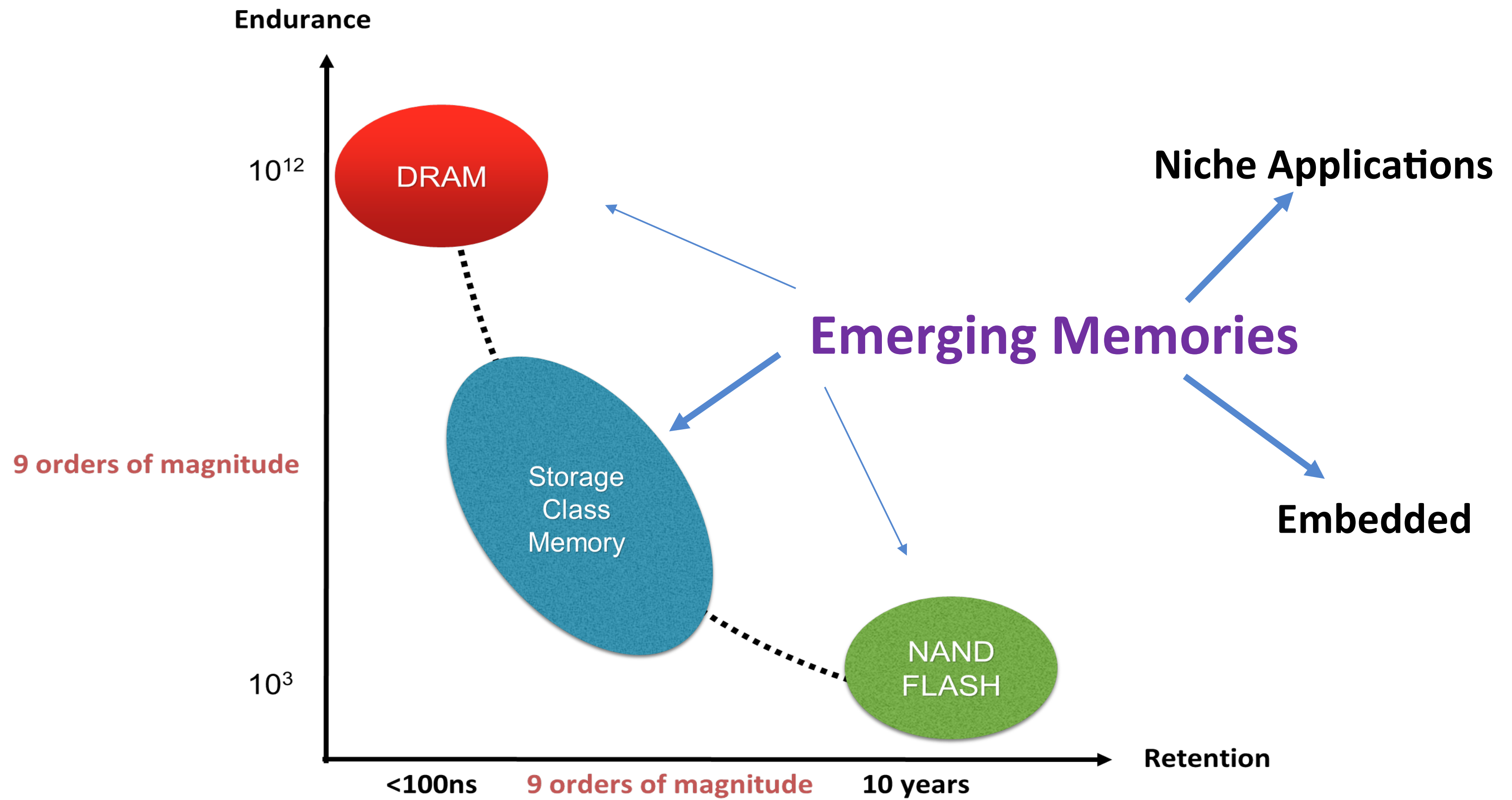
Challenges	Notes
Structural	Insufficient room to fit the capacitor, bit-line, word-line, two contacts, isolation, etc.
Capacitor	Reduction in cell capacitance will cause reduced refresh time, increased ECC and power, etc.
Transistor	Scaling of the access transistor while achieving high drive (>10μA) and very low leakage (<1fA)
Speed	Increased bit-line and word-line resistance along with increased parasitic capacitance between all nodes make it difficult to maintain array speed

Desu-2017

NAND Scaling Limitations

Challenges	Notes
Effective Scaling	3-D scaling will continue in the short term with an increasing number of layers stacked vertically with limited horizontal scaling. Limitations??
ECC	NAND is pushing the limits of ECC already correcting one bit in a 100
Defect Density	Un-repairable defect density may become a limitation
Cost	Scaling may be technically feasible, but the cost to achieve scaling for a future nodes may not result in a decreased cost per bit

Semiconductor Memories



Desu-2017

Recent Success of Emerging Memories

Storage Class Memories

3D-Xpoint based Optane Memories and Storage by Intel

You can buy 3D-Xpoint Memory from FRYS!

FRYS.COM

keyword, model#, or frys.com# GO

Cool Stuff We Sell | Things We Do for You | myFrys | Welcome Guest! | Cart

Frys» PC Components» CPU/Processor» #9192408

Intel® Optane™ Memory Series 32GB M.2 80mm

Price: $79.99

Add To Cart

Frys.com #9192408 Manufacturer: Intel
UPC #735858337090 Model #8X0504
Detailed Description | Warranty Info

Performance

- Sequential Read (up to) : 1350 MB/s
- Sequential Write (up to) : 290 MB/s
- Random Read (8GB Span) (up to) : 240000 IOPS
- Random Read (100% Span) : 240000 IOPS
- Random Write (8GB Span) (up to) : 65000 IOPS
- Random Write (100% Span) : 65000 IOPS
- Latency - Read : 7 µs
- Latency - Write : 18 µs
- Power - Active : 3.5 Watts
- Power - Idle : 1 Watt

Overall System Performance	28%
Hard Drive Access	14%
Task Responsiveness	2%

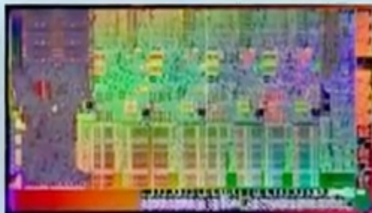
Desu-2017

3D XPOINT™ MEMORY MEDIA

Breaks the memory/storage barrier

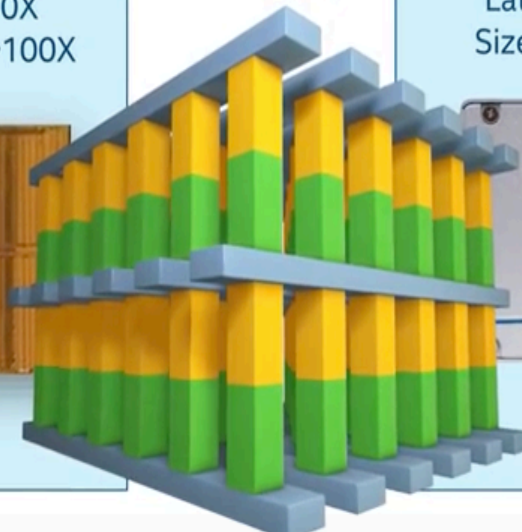
MEMORY + **STORAGE**

SRAM
Latency: 1X
Size of Data: 1X



DRAM
Latency: ~10X
Size of Data: ~100X

3D XPoint™
Latency: ~100X
Size of Data: ~1,000X

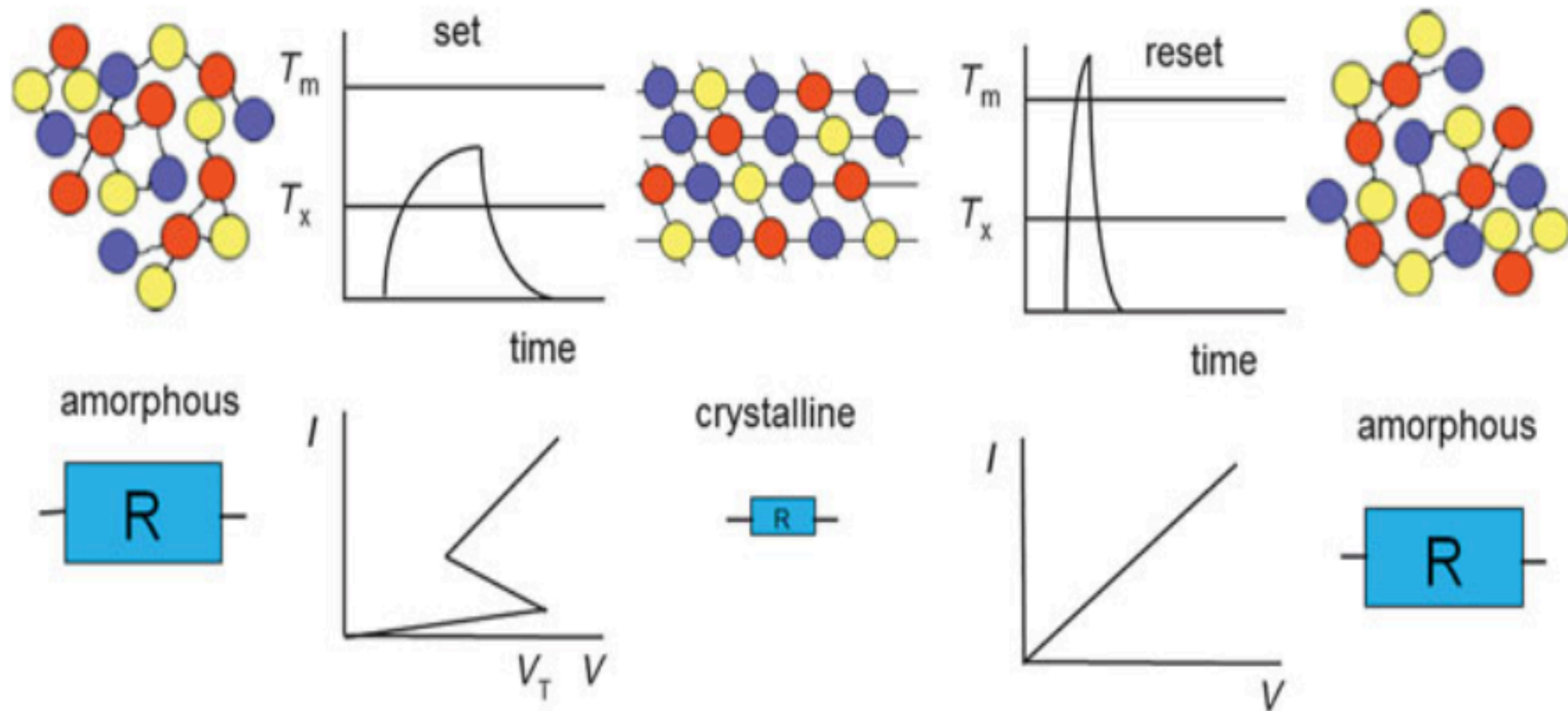


NAND SSD
Latency: ~100,000X
Size of Data: ~1,000X

HDD
Latency: ~10 Million X
Size of Data: ~10,000X

Desu-2017

Principle of Phase Change Memory (PCM)



Desu-2017

3D-Xpoint Memory and Storage

Intel® Optane™ Solid State Drives for Consumers

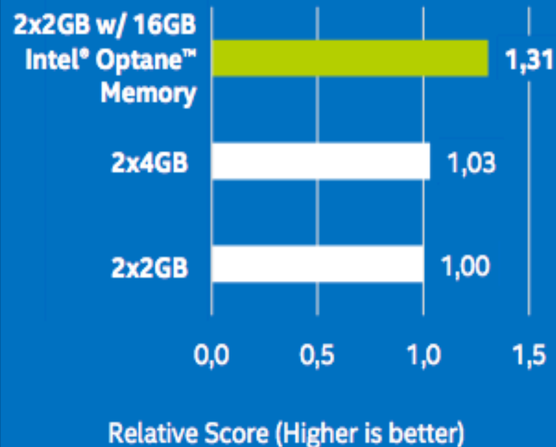
Intel® Optane™ Technology Data Center Solutions

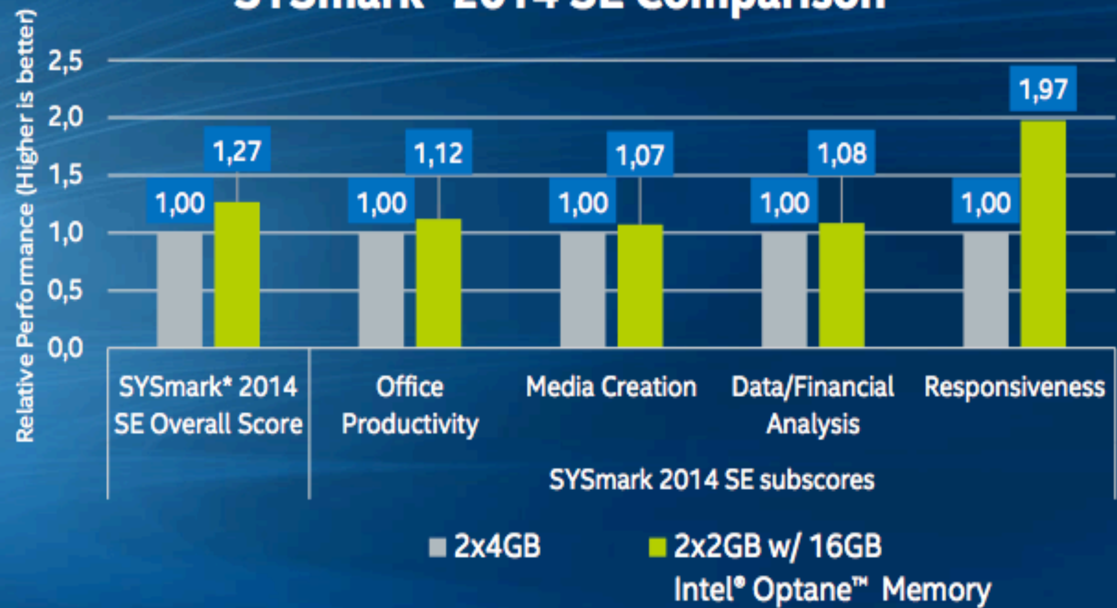
Desu-2017

INTEL® OPTANE™ MEMORY OFFERS BETTER END USER VALUE

16GB Intel® Optane™ memory + 1TB HDD 4GB DDR delivers better responsiveness than 1TB HDD 8GB DDR

Intel® Optane™ Memory Delivers Visible Benefits

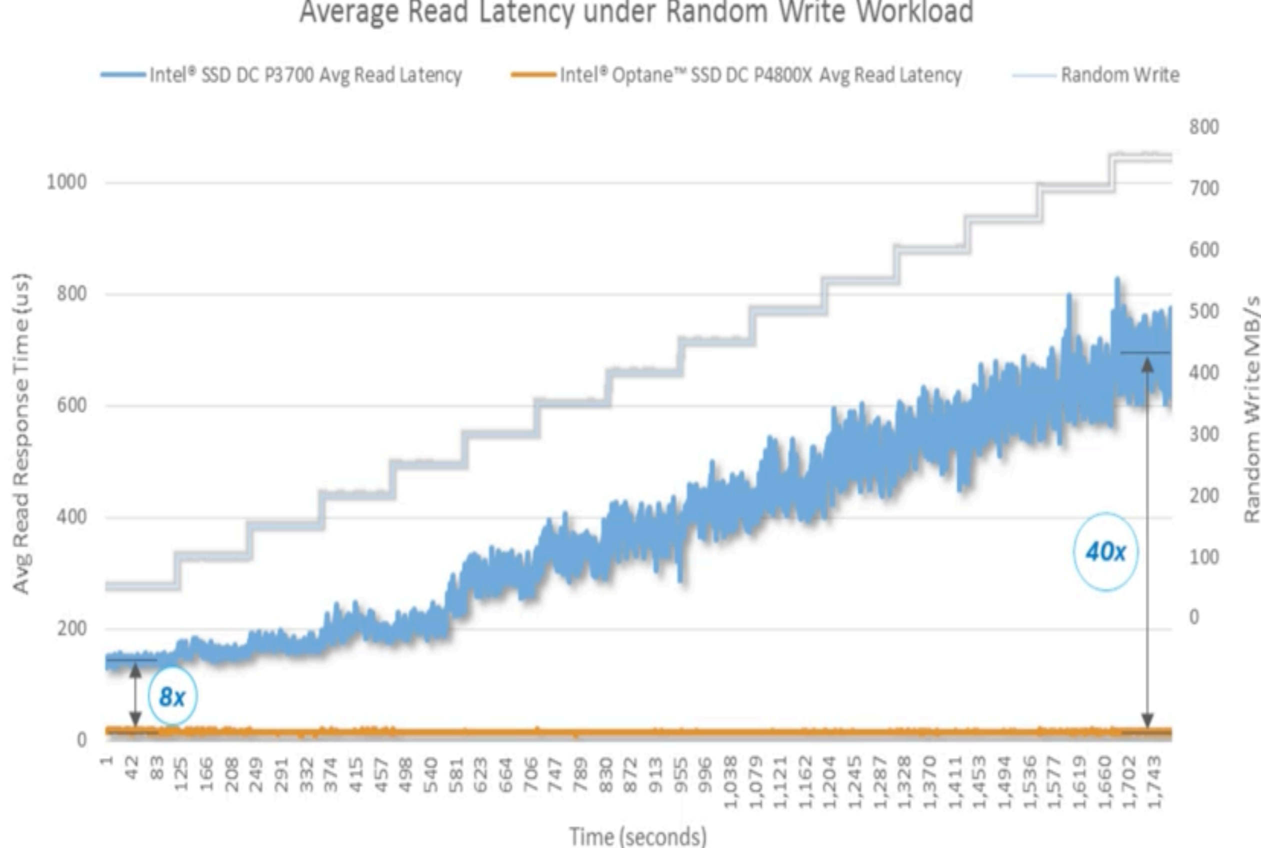




All testing done internally by Intel.
See Appendix I for System Configuration and testing procedures.
Note: Some applications and most games may require 8GB of DRAM memory for loading with or without Intel® Optane™ memory.

Desu-2017

Responsive Under Load



+

- up to **40X faster response time** under workload[1]
- Consistently **amazing response time under load**

Desu-2017

Recent Success of Emerging Memories contd.

Niche Applications

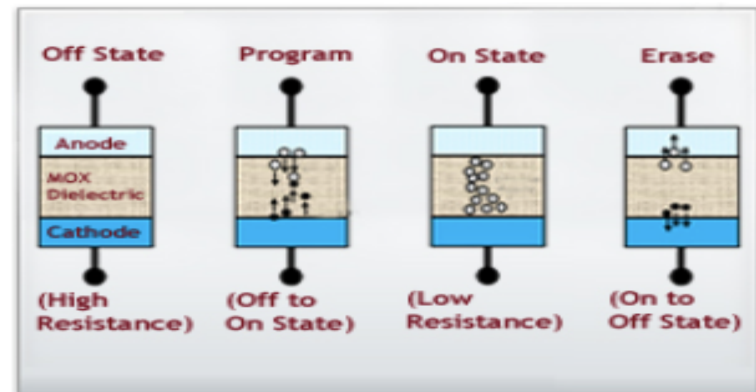
- EverspinTM has used its Toggle-MRAM in high cycle, high reliability applications and is pushing upwards in density with STTRAM
- AdestoTM has used CBRAM to undercut the cost of EEPROMs and is pushing its technology into NOR densities
- PanasonicTM has effectively used metal- oxide RRAM as an embedded memory

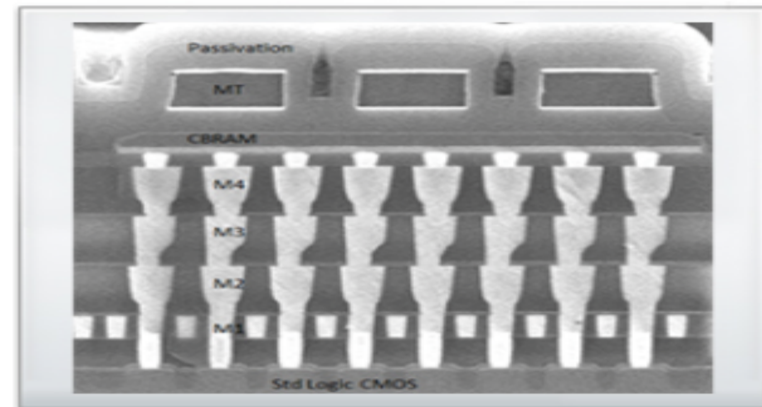
Everspin announced it is sampling a Gbit MRAM chip

- **Gbit chip uses a DDR4 interface and is made in Globalfoundries' 28nm process**

Adesto Introduces Ultra Low Power EEPROM Mavriq™ DS Memory Family up to 512Kbit density

- **It is built on Adesto's RRAM technology, known as Conductive Bridging RAM (CBRAM®)**
- **It performs read and write operations with 4x less power than competitive solutions, and, in ultra-deep power down mode, uses as much as 50x less power.**
- **Provides over 100,000 write cycle endurance across the full temperature and voltage range**



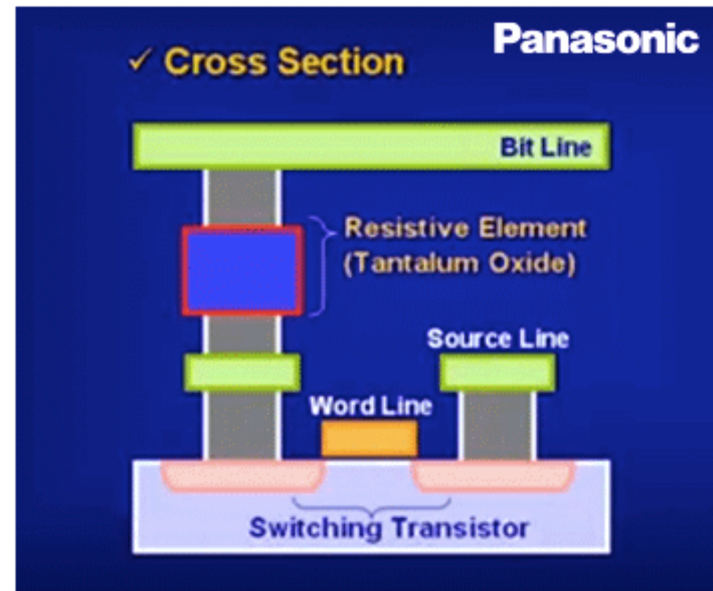


Desu-2017

Panasonic and UMC Partner for 40nm ReRAM Process Platform

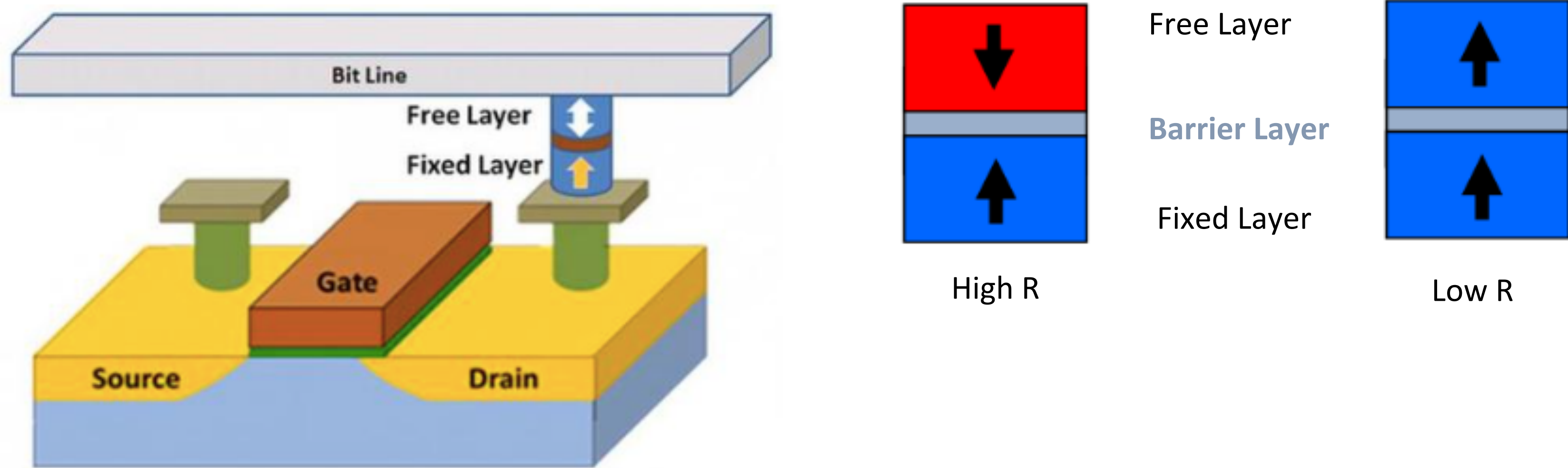
- **Enable the integration of PSCS's 40nm ReRAM with UMC's CMOS process to achieve a ReRAM technology platform that incorporates embedded memories in place of flash**

- **Expected to ship product samples based on UMC's 40nm process in 2018**



STT-MRAM

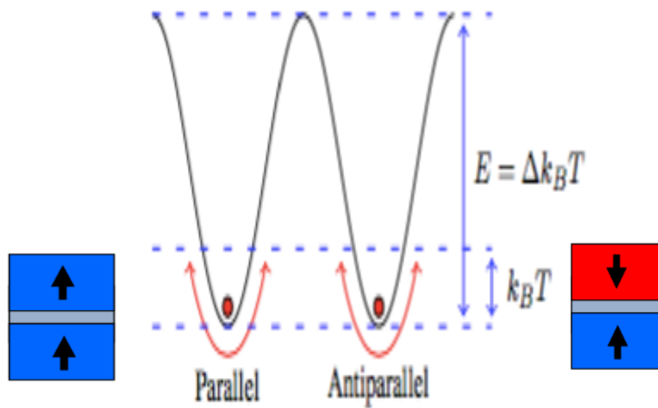
The difference in resistance between Parallel and Anti-Parallel spin arrangements of MTJ is the key for STT-MRAM



- Read operation by probing the resistance of the device at low voltage bias

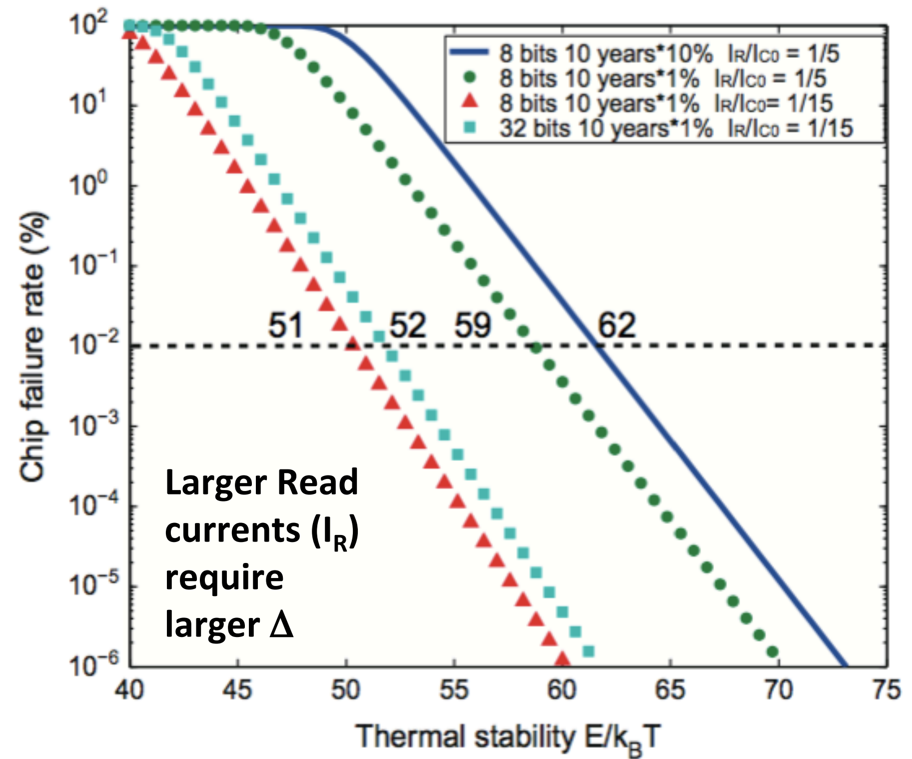
Desu-2017

Thermal Stability Factor (Δ) is the key metric of an MTJ



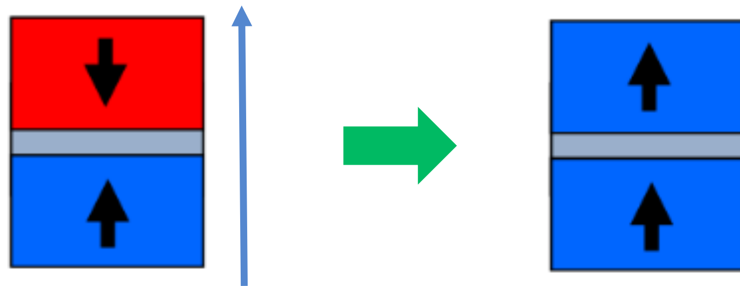
High thermal stability factor (Δ) is required to reduce the erroneous sensing rate

$$F_{chip} = 1 - \exp\left[-N\frac{\tau}{\tau_0}\exp\left(-\Delta\left(1-\frac{I_R}{I_{CO}}\right)\right)\right]$$



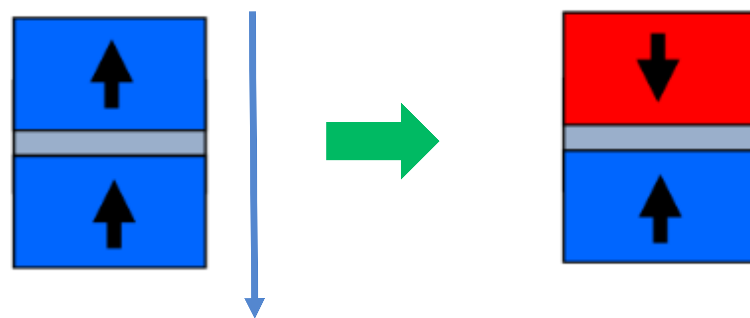
Desu-2017

Write operation consists of transfer of spin-angular momentum from polarized conduction electrons to electrode magnetization



As electrons move through the fixed layer, their spin is polarized to match the fixed layer. When a sufficient density of polarized electron flow is achieved, the free layer changes state to align with the fixed layer

Current Direction



In this switching direction, electrons with spin opposite the fixed layer are reflected back to the free layer causing the free layer to switch to the state opposite to the fixed layer

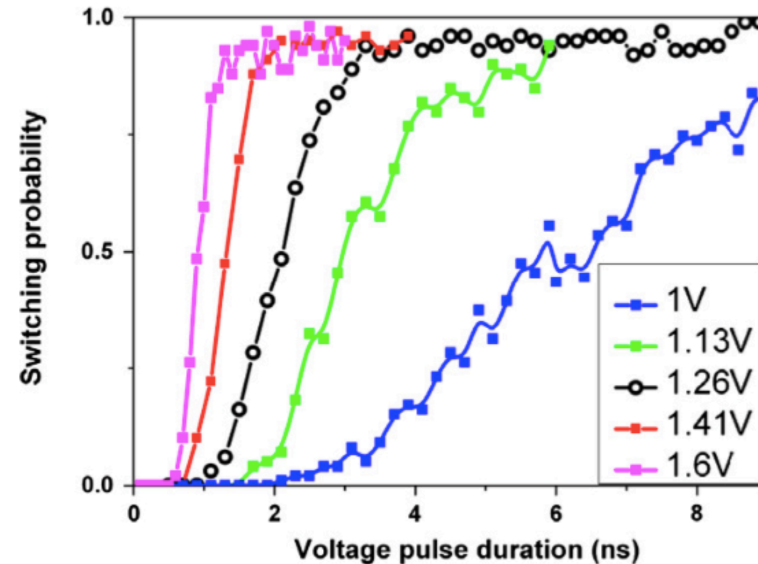
Desu-2017

Trade-offs of STT-MRAM Writing

- **Non-Deterministic Write**

- **STT write process is inherently stochastic** *(sigmoidal distribution with very long tail)*
- **The stochasticity of switching time is temporal** *(leading to variation in transition time for a single cell)*

- STT vanishes for parallel alignment
- Switching time inversely proportional to angle between the layers
- Thermal fluctuations provide initial 'kick'

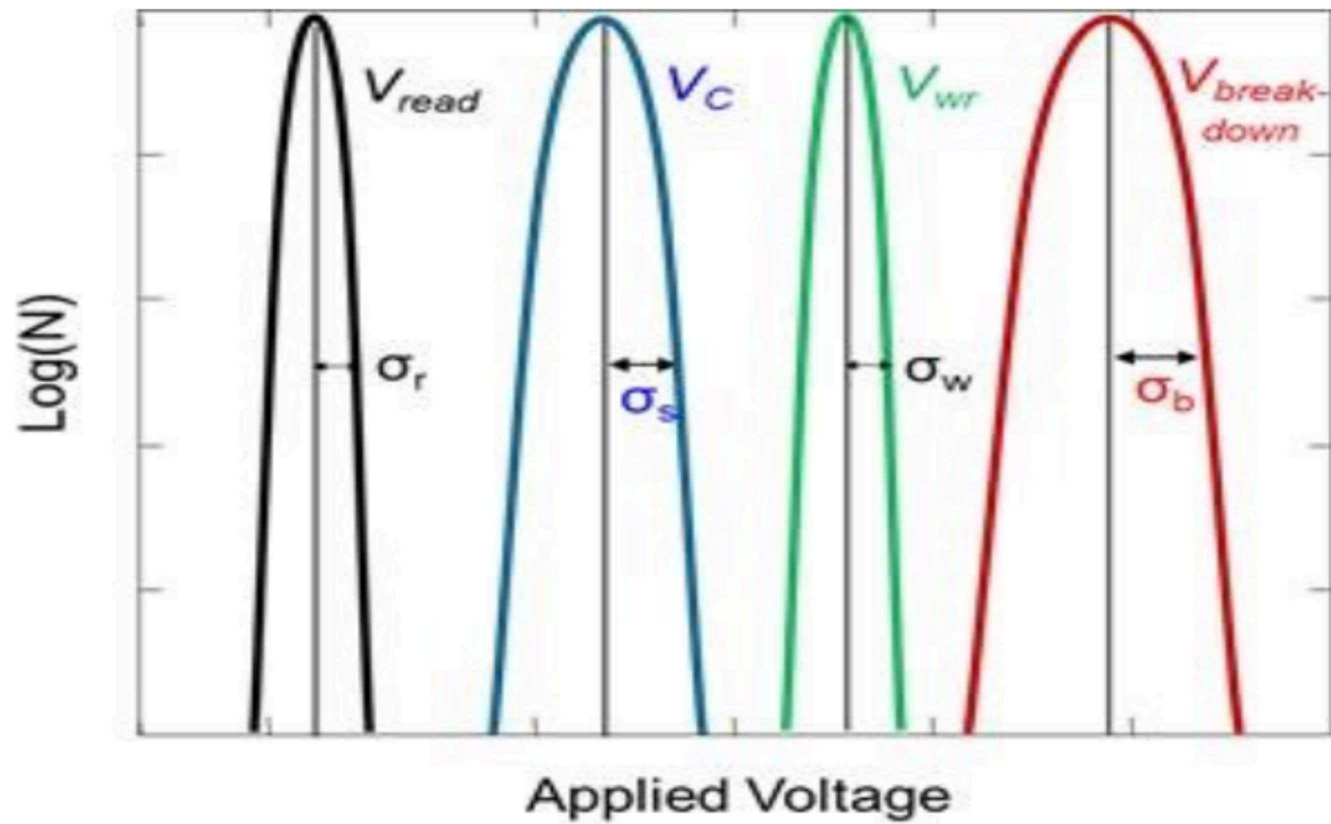


STT stochastic switching behavior

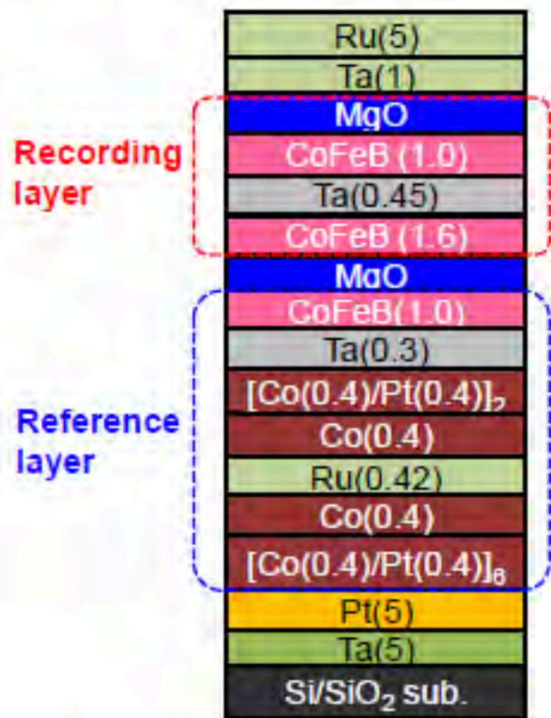
- Increasing the write current value I_{WR} or driver pulse duration are the most efficient methods to avoid the writing failures
- But lead to significant power, speed, surface overhead storage, and could drive the breakdown or damage of oxide barrier

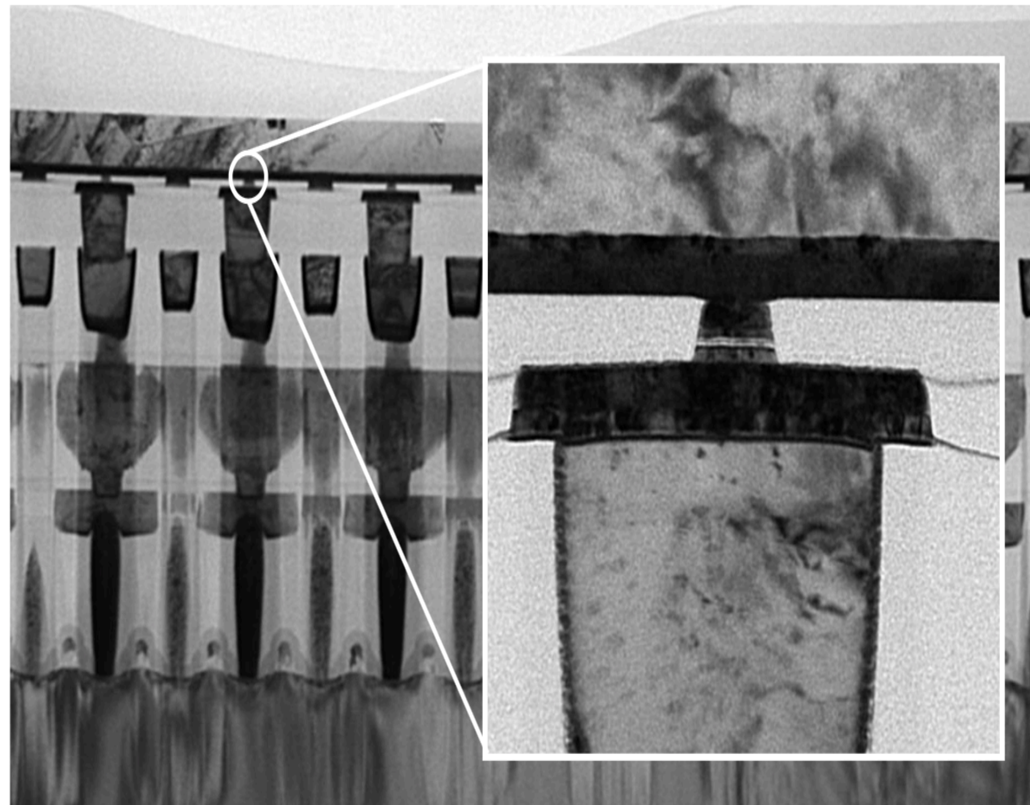
Desu-2017

Major Challenge for STTRAM



MTJ stack is about 20 nm thick, can be easily integrated into CMOS backend process: ~3 additional masks





Desu-2017

STT-MRAM is well suited for Embedded

- Combines non-volatility, high speed, and infinite endurance
- Higher energy efficiency (mobile and IoT applications)
- Opportunity for eMRAM as Last Level Cache
- Lower Cost

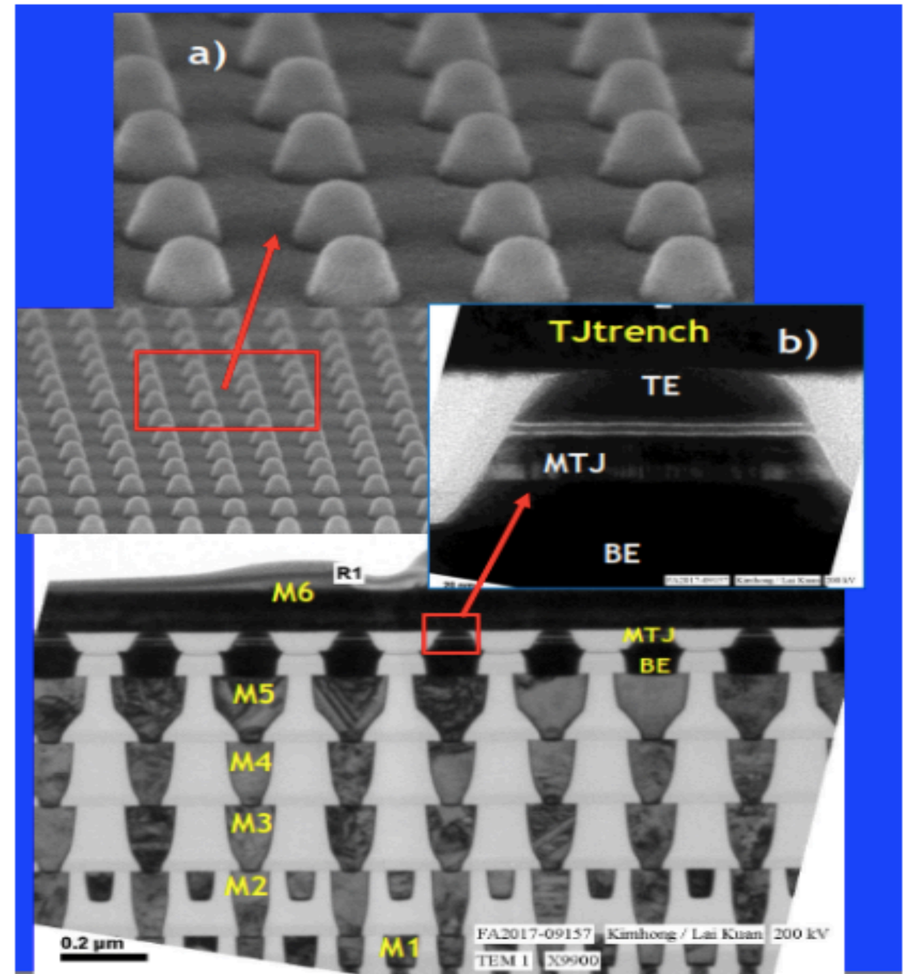
	eMRAM	eFlash	eDRAM	SRAM
Cell size (F^2)	30-50	30-50	30-90	100-300
Added mask layers	2-3	10-12	4-6	0

Desu-2017

Global Foundries has plans to deploy Everspin's
p-MTJ STT-MRAM as an embedded 22nm memory - as part of GF's 22FDX platform.

Claims for 22FDX-eMRAM:

- 1,000x faster write speeds
- 1,000x higher endurance
- Scalable beyond 22nm (FinFET or FDX)

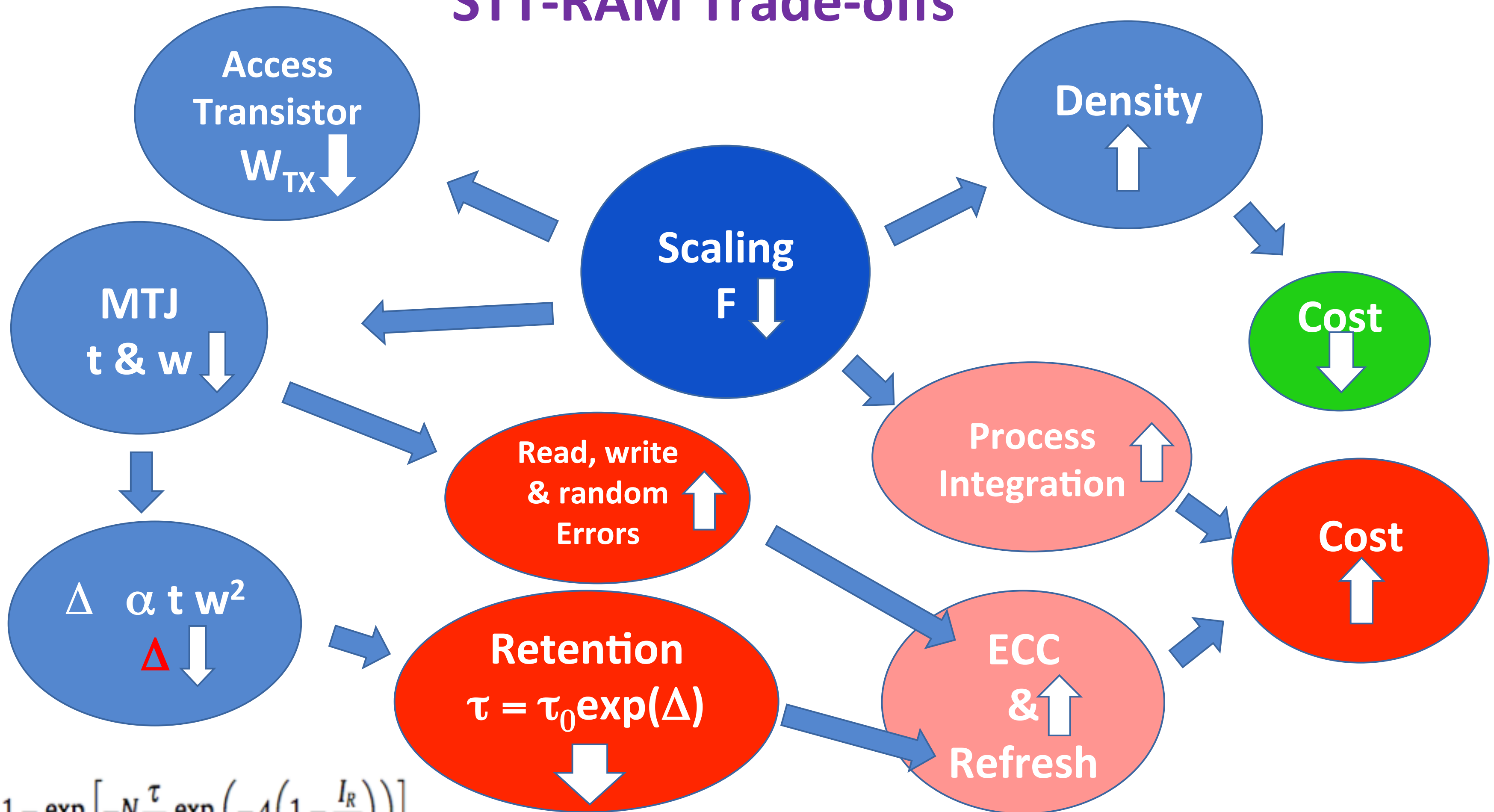


Desu-2017

[Samsung reaffirms 2018 target for STT-MRAM mass production]

[TSMC to start eMRAM production in 2018]

Desu-2017

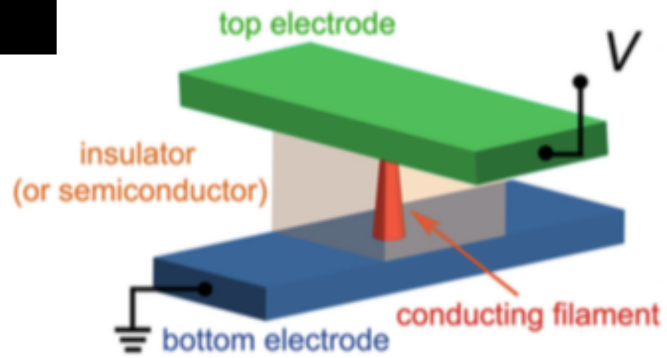

STT-RAM Trade-offs
Access Transistor W_{TX}
Scaling F
Density
MTJ t & w
Cost
Read, write & random Errors
Process Integration
$\Delta \; \alpha \; t \, w^2$
Δ
Retention
$\tau = \tau_0 \exp(\Delta)$
ECC & Refresh
Cost
$F_{chip} = 1 - \exp\left[-N\frac{\tau}{\tau_0}\exp\left(-\Delta\left(1-\frac{I_R}{I_{CO}}\right)\right)\right]$
Desu-2017

STT-MRAM may not replace DRAM in the near-future

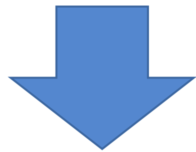
- **Cost is directly related to density & cell/chip size**
- **Current available scales with transistor size**
- **Standalone DRAM : GB chips, cell size ~4F2**
  F smallest feature at technology node (16/14nm,...)
- **MTJ < 20 nm**
- **Write current < 20 μA**
- **TMR ~ 300%**

Kent & Worledge, Nature Nano (2015)

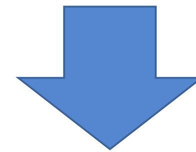
Filament RRAM



Filament ReRAM



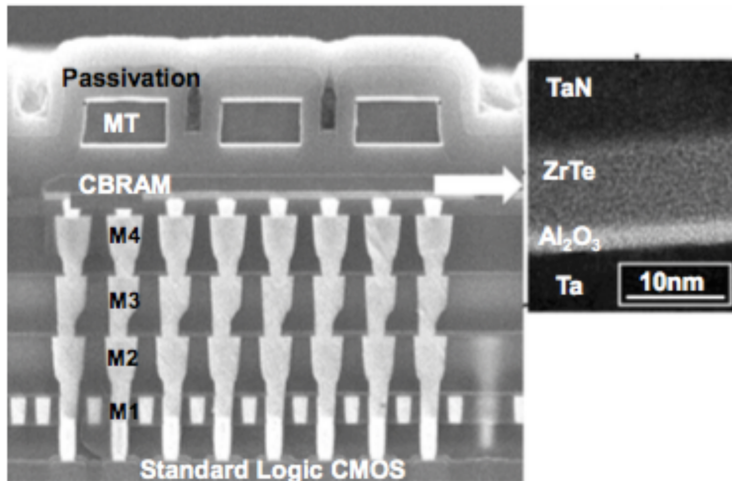
Oxy-ReRAM



CBRAM

Desu-2017

Adesto 512 Kbit EEPROM-compatible CBRAM



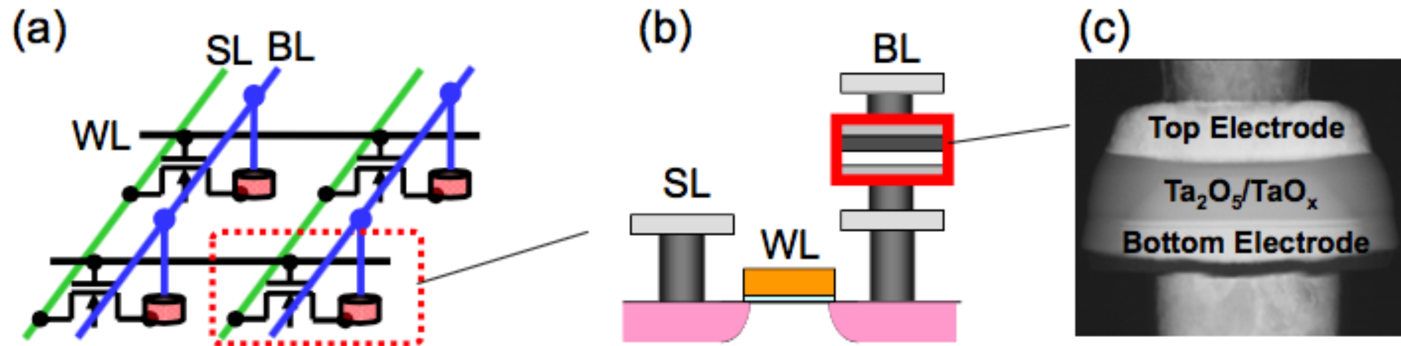
Parameter	Adesto Moneta	Standard EEPROM
Core Supply Voltage	0.97–1.03V	Not Applicable
I/O Supply Voltage	1.65–2.75V	Not Applicable
Single Supply Voltage	Not Applicable	1.7–5.5V
Read Power (500Kb/s)	10μW	1250μW
Lowest Power-Down Mode	.05μW	.25μW
Clock Frequency (Max)	1MHz	20MHz
Operating Temp Range	–40 to +85°C	–40 to +85°C
Write Supply Voltage	3.6–4.4V	1.7–5.5V
Write Power (10Kb/s)	7.5μW	375μW

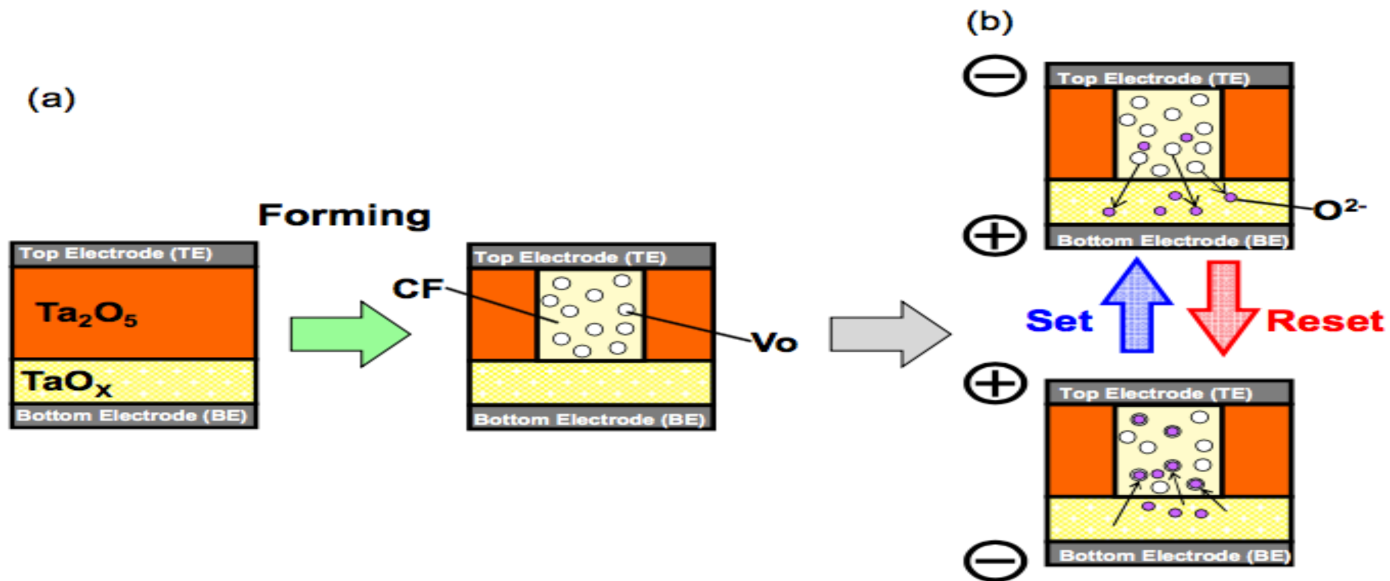
- **Single supply voltage: 1.65V - 3.6V**
- **1.6 MHz maximum clock rate for normal read**
- **20 MHz maximum clock rate for fast read**
- **Byte Write consuming 50 nJ**
- **0.25 mA Read current; 1 mA Write current**
- **Byte Write within 25 μs**
- **Data Retention: 10 years**
- **Endurance: 10,000 Write Cycles**
- **Unlimited Read Cycles**

Te---Metalloid Filament!!

Desu-2017

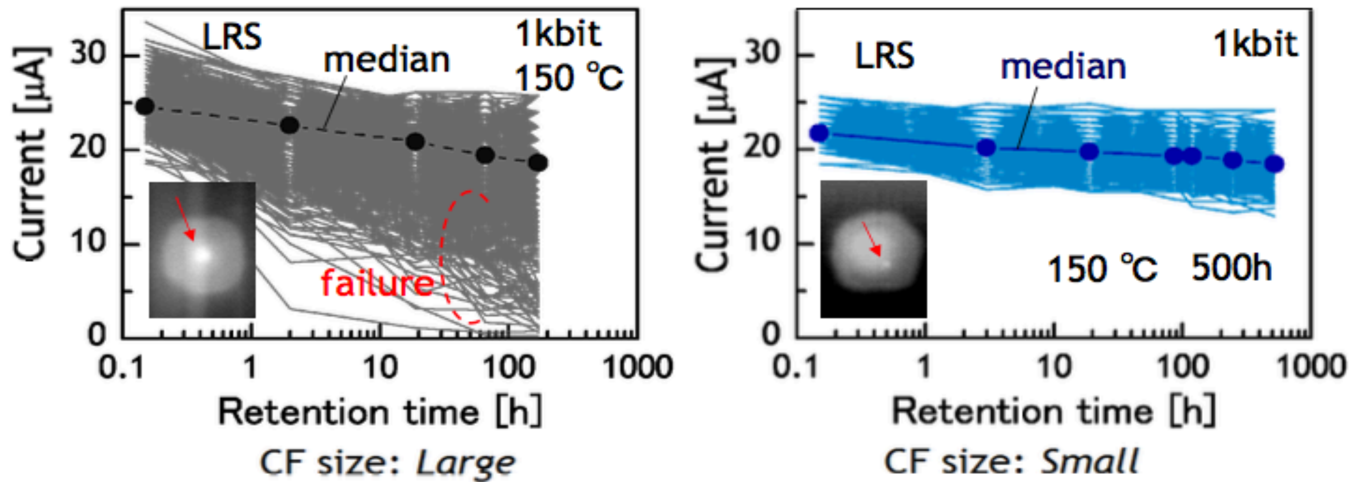
Panasonic ReRAM Technology-TaOx based

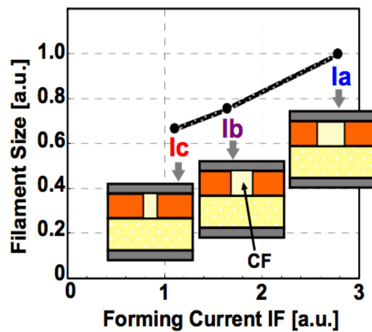


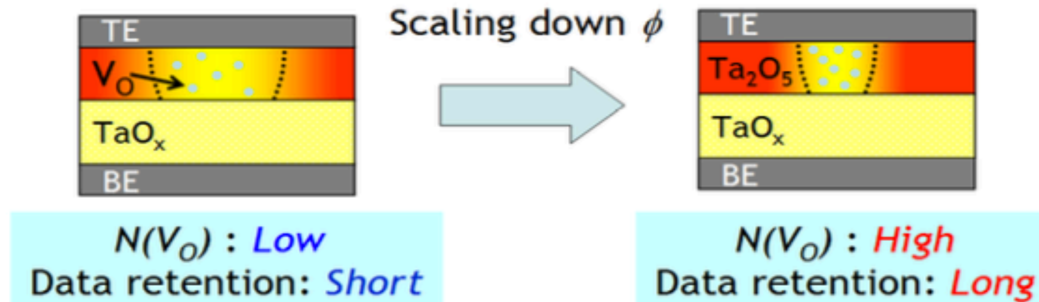


Desu-2017

Relationships between CF characteristics and density of Vo and retention characteristics

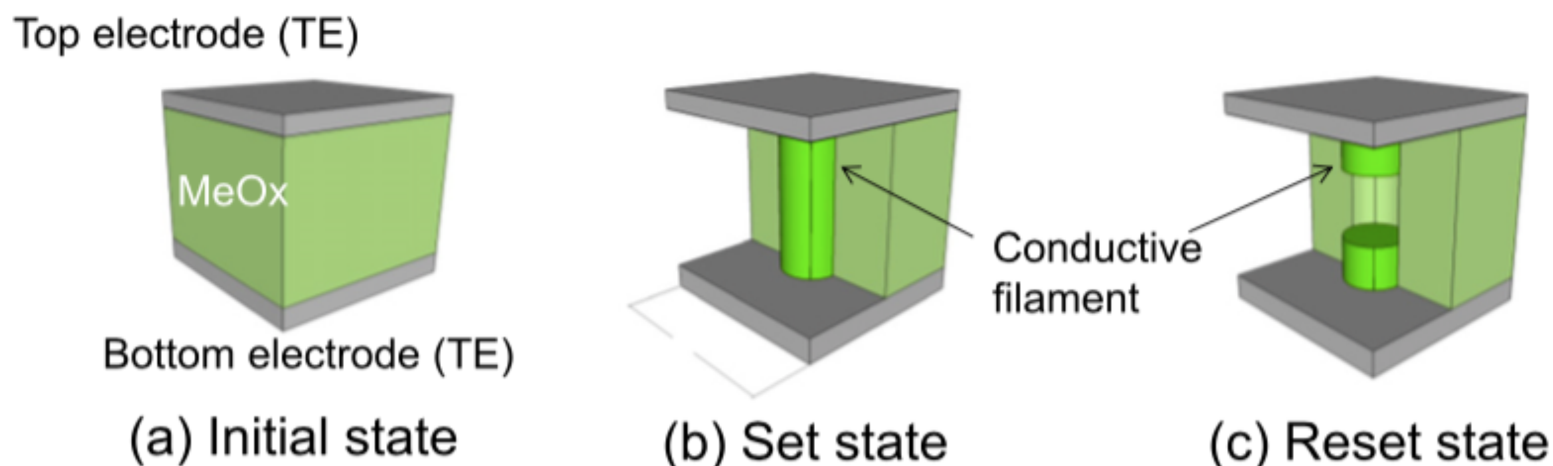




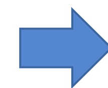


Desu-2017

In addition to the process variations, Filament formation, rupture and regrowth are stochastic processes!!



- **Local high fields can change due to change in nature of defects and their concentrations (> 1 MV/cm)**
- **Local power dissipation can alter the local temperatures (~ 1 TW/cm^3)**
- **Local electronic current densities can vary (> 10^6 A/cm^2)**
- **Significant local variations in ionic current densities**

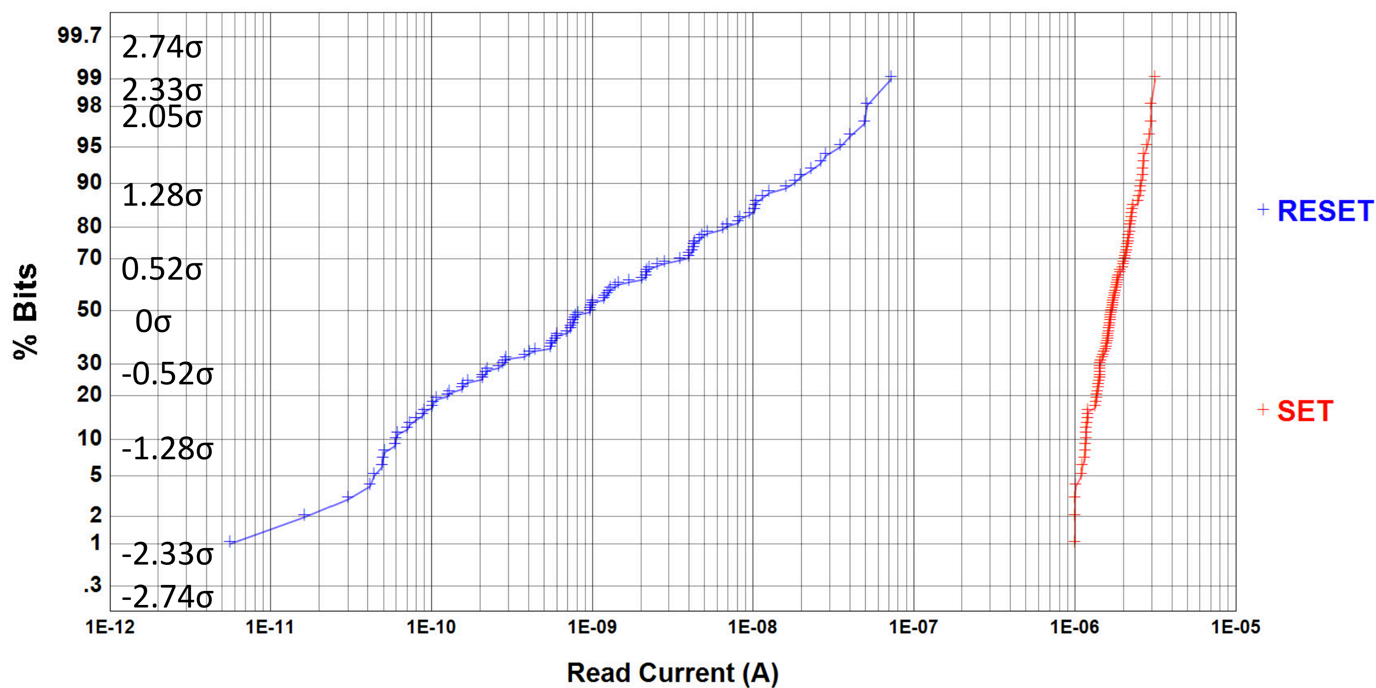


Verify & ECC

Extensive optimization needed to engineer the tail bits

Desu-2017

Filamentary ReRAM Distribution Example



19.7: A 16Gb ReRAM with 200MB/s Write and 1GB/s Read in 27nm Technology

G. Atwood et al., ISSCC 2014

Filamentary read current fluctuations

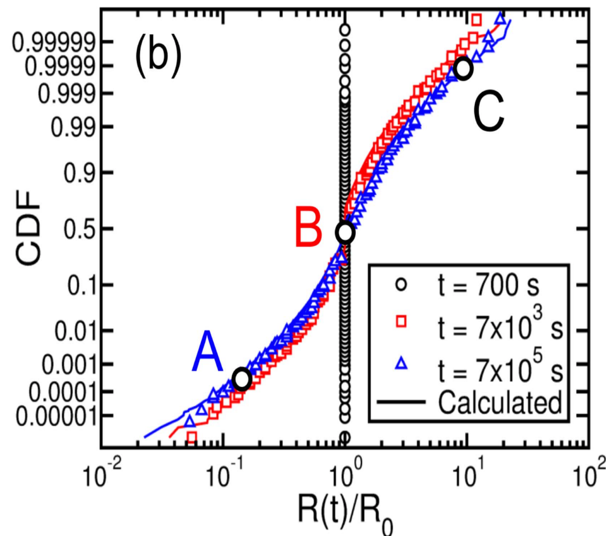


Fig. 1. Measured and calculated distributions of (a) R and (b) $R(t)/R_0$ for increasing times 700, 7×10^3, and 7×10^5 s. Note the increasing tails in (b) due to current fluctuations. Selected cells A, B, and C are also shown in their final position within the distributions at $t = 7 \times 10^5$ s.

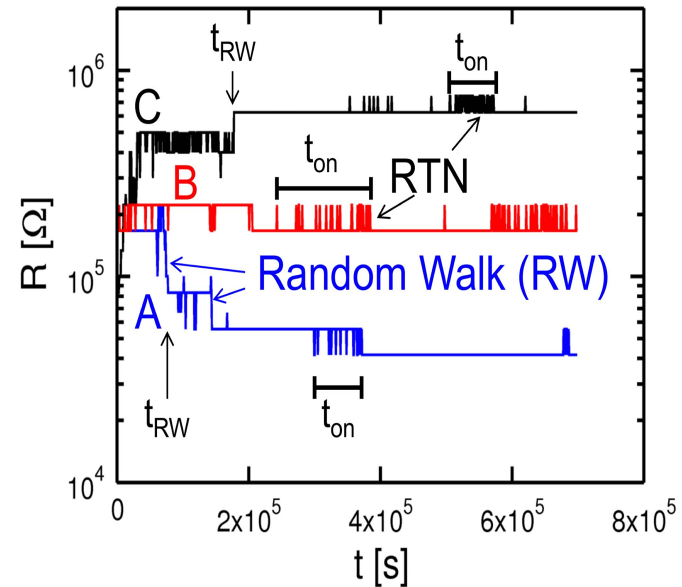
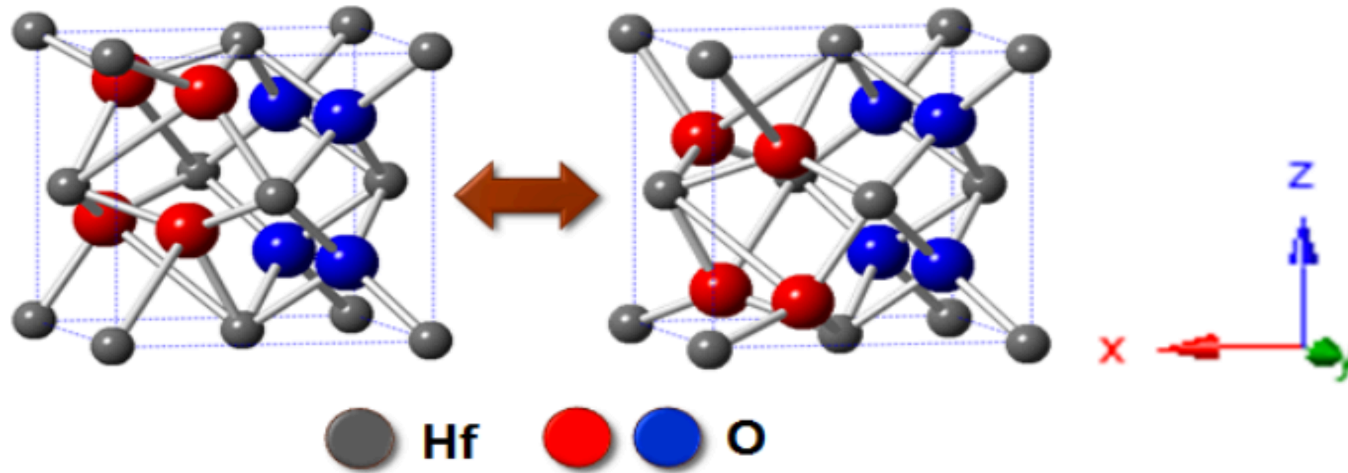


Fig. 2. Measured R as a function of time for cells A, B, and C in Fig. 1, showing RW occurring at t_{RW} and interrupted RTN lasting for time t_{ON}.

Source: IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 62, NO. 11, NOVEMBER 2015
Noise-Induced Resistance Broadening in Resistive Switching Memory—Part II: Array Statistics
Prof. Daniele Ielmini, Senior Member, IEEE, Politecnico di Milano, Milan, Italy

FeFET

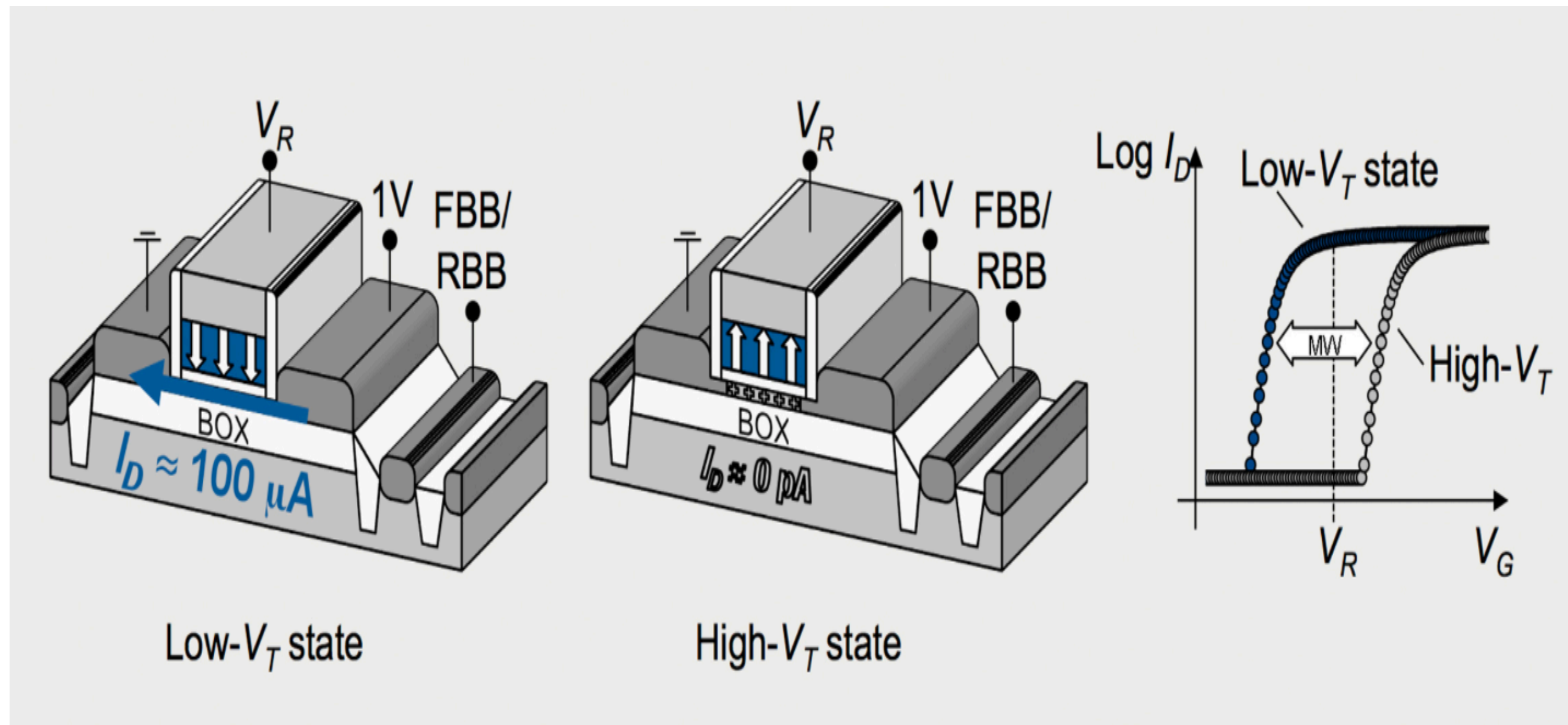
Ferroelectric Switching in thin HfO2 Films



- The stable ferroelectric phase in HfO2, the orthorhombic phase $Pca2_1$.
- The movement of four active oxygen atoms mainly contributes to the ferroelectric switching.
- The crystal distortion is driven by symmetry conserving phonon modes

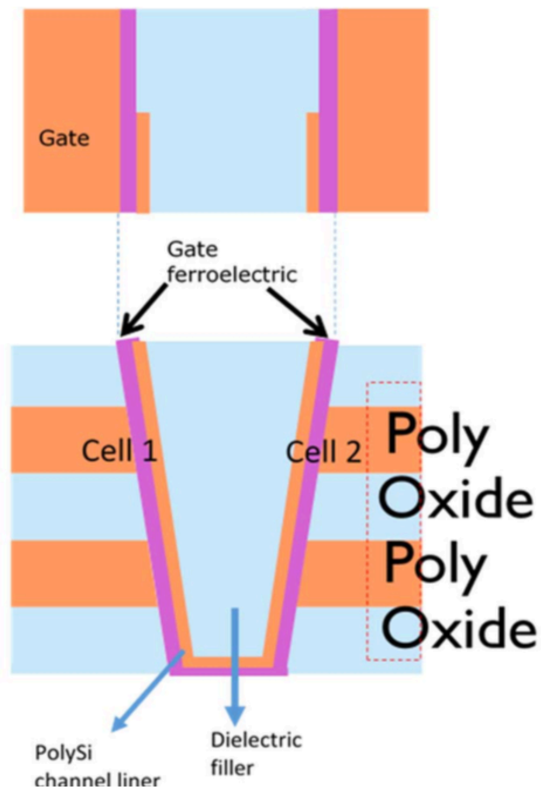
Desu-2017

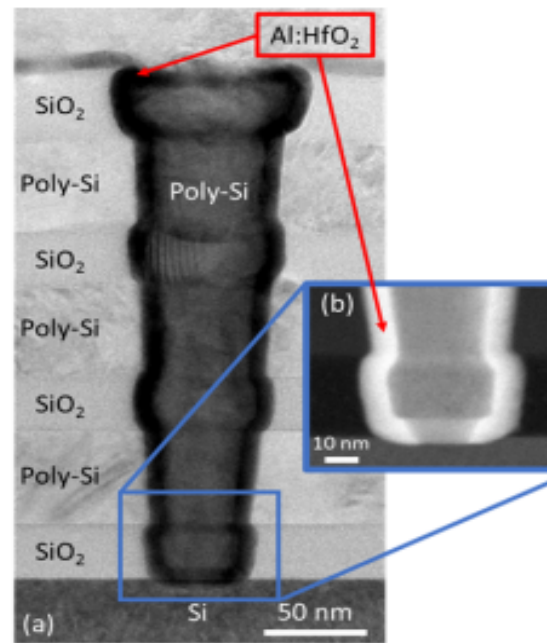
Typical FeFET structure



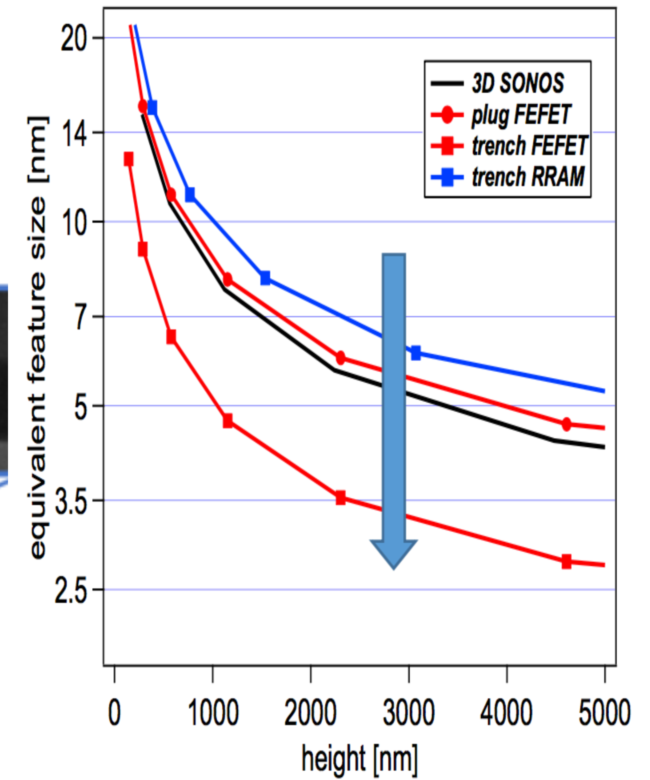
Desu-2017

Trench-based 3D FeFET memory



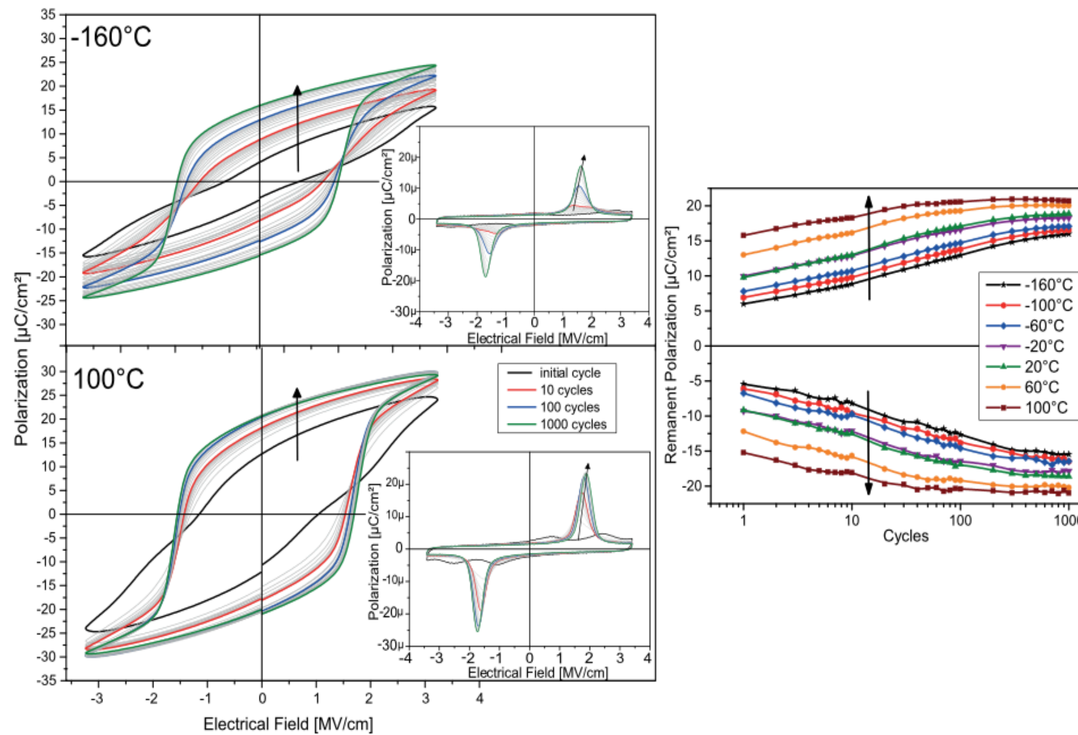


Imec's FeFET (Source: Imec)



Desu-2017

Oxygen vacancies movement during wake-up in Ferroelectric Hafnium Oxide



- **The duration of the applied electrical field, not the amount of cycles, is essential for the wake-up**
- **A strong temperature activation of the wake-up**
- **Attributed to ion rearrangement during cycling**
- **Resistive valence change mechanism switching can be observed coincident with ferroelectric switching depending on electrodes**

S. Starschich, et al, Appl. Phys. Lett. 108, 032903 (2016)

Challenges and Opportunities of EM

Switching Speed Limitations

Emerging Memories	Fundamentals	Switching Time	Tradeoffs	Bandwidth
PCM	Temperature ramp down for crystallization	100–400nS	Separation of states	9MB/s
STTRAM	Stochastic switching	10-50nS	TDDB Bit Error Rate	2.66GB/s
Filament RRAM	Ion migration	10–100nS	Retention Separation of states	200MB/s

Desu-2017

Multilevel Challenges

EM	Multilevel Capability
PCM	Drift of the high resistance state
STTRAM	Lack of voltage window Very low difference between states
Filament RRAM	Data retention of partially formed filaments for intermediate states Stochastic behavior

Desu-2017

Endurance Limitations

EM	Endurance Failure Mode
PCM	Atom segregation (local stoichiometry variation) Physical separation of electrode due to volume change
STTRAM	TDDB
Filament RRAM	TDDB Over-accumulation of atoms in filaments

Read and Write Disturb

EM	Read Disturb	Write Disturb
PCM	Crystallization of the amorphous state	Thermal disturb
STTRAM	Increased probability of stochastic switching	Oxide Breakdown
Filament RRAM	Dissolution of the filament	Excessive filament formation

Emerging Memory Opportunities

- **SCM: 3D-Xpoint**
- **Niche Applications: CBRAM for EEPROMs and NOR**
- **Embedded Memories: STTRAM and Filament RRAM**

- **No EM technology, at this point, that is a serious contender against NAND: lack of MLC, 3-D, etc**
- **No EM technology that can compete with the lower energy and high bandwidth of DRAM, which are directly a result of the DRAM cells low operating energy**

Thanks

Questions and/or Comments

ReRAM Comparison

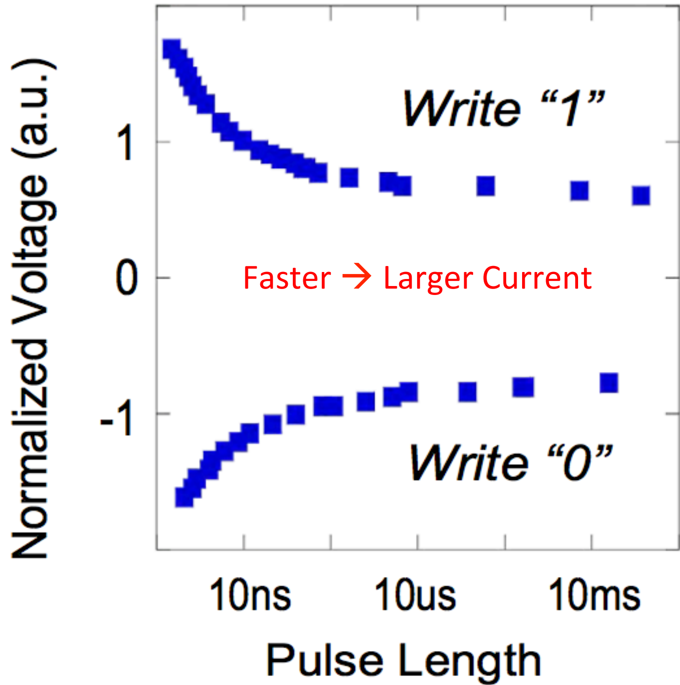
	Filamentary	Interface
Switching mechanism	• Create/eliminate "atomic" wire • via oxygen vacancy or metal injection) in switching film	• Bulk resistance change near oxygen donor-electrode interface • Oxygen donor-electrode oxygen exchange
Switching location	Point based	Area based
Maximum cell current	• Compliance current dependent • Cell area ***independent***	• Initial (OOF) • Cell area ***dependent***

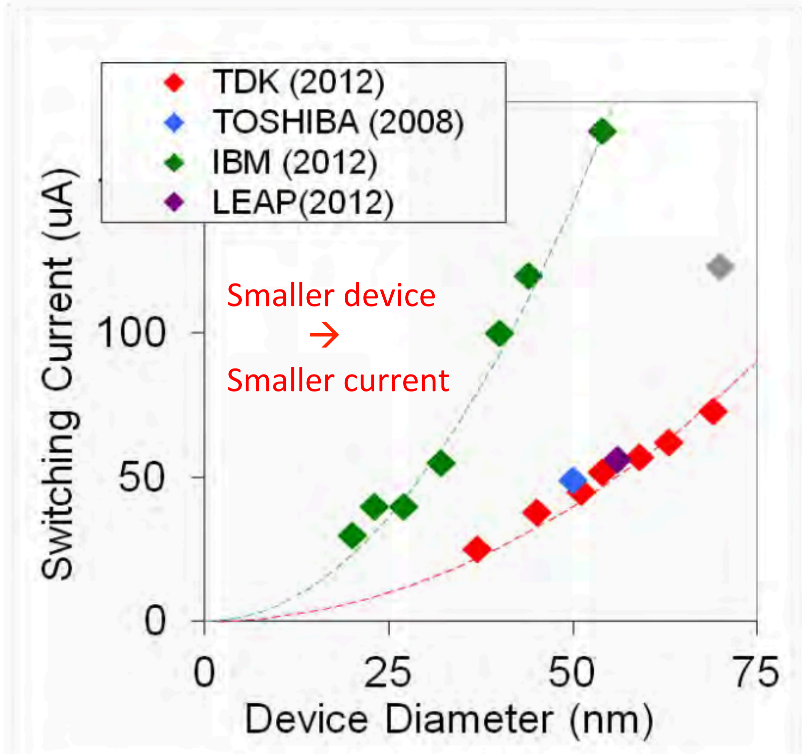
Resistance fluctuation	• High probability (esp. $10K\Omega < R_{cell} < 100K\Omega$)
	• Function of compliance current
	  ◦ Low = high fluctuation
	  ◦ High = lower fluctuation; high current density
		• Low probability
		• Area based conduction

Retention	Diffusion limited	Diffusion limited

Desu-2017

STT switching current inversely proportional to pulse width but scales with area



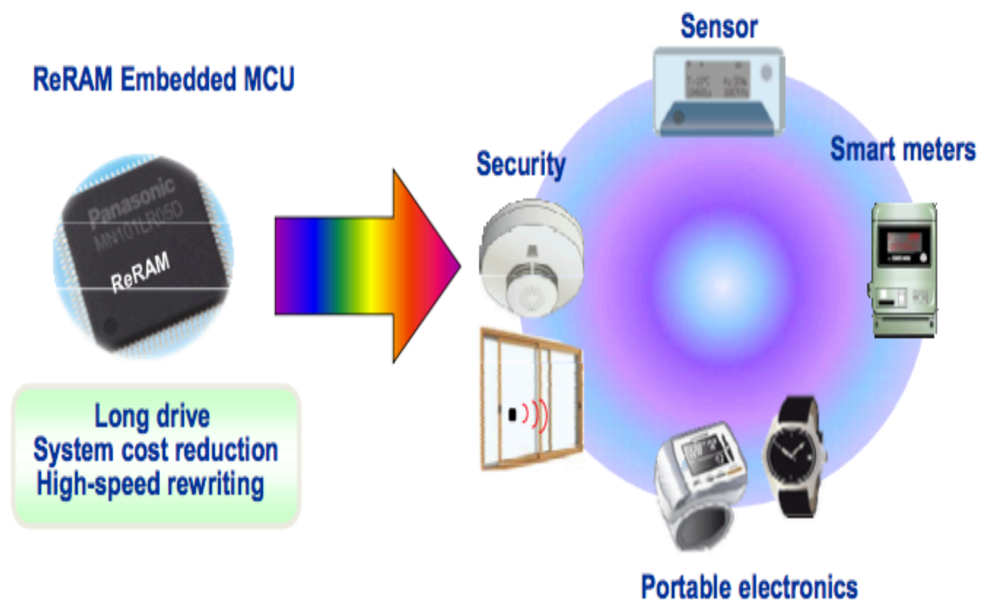


Desu-2017

Features and Applications of Panasonic ReRAM (0.18 μm)

Density	64KB
Power Supply	1.1～3.6V
Frequency	10MHz (1.8~3.6V) 1MHz (1.3~3.6V) 40kHz (1.1~3.6V)
Current	2.1mA@10MHz 0.22mA@1MHz 5.6μA@32.768kHz
Endurance	1K (Program Area) 100K (Data Area)
Retention	85℃, >10years

EMBEDDED MCU (MN101LR05D)

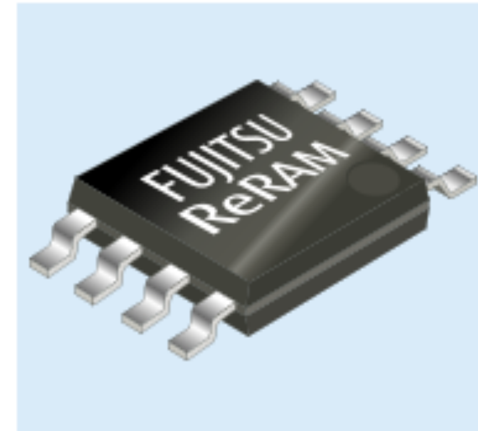


Desu-2017

Fujitsu Semiconductor Launches World's Largest Density 4 Mbit ReRAM Product for Mass Production

Key Specifications

- Product Part Number : MB85AS4MT
- Memory Density (configuration) : 4 Mbit (512K words x 8 bits)
- Interface : Serial peripheral interface (SPI)
- Operating power supply voltage : 1.65V – 3.6V
- Low power consumption :
  - Read operating current : 0.2mA (at 5MHz)
  - Write operating current : 1.3mA (during write cycle time)
  - Standby current :10μA
  - Sleep current :2μA
- Guaranteed write cycles : 1.2 million cycles
- Guaranteed read cycles : Unlimited
- Write cycle time (256 byte page) : 16ms (with 100% data inversion)
- Data retention : 10 years (up to 85℃)
- Package : 209mil 8-pin SOP



EVERSPIN TECHNOLOGIES The MRAM Company

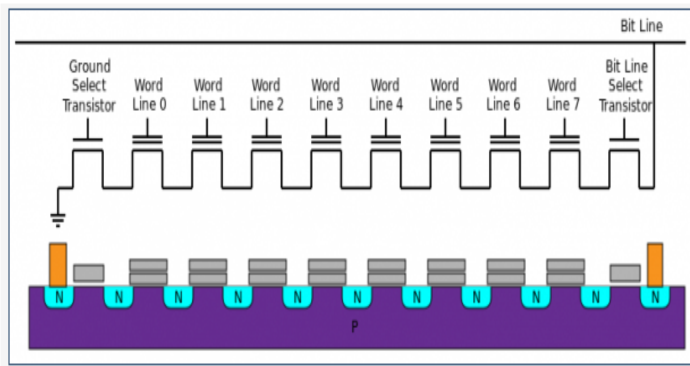
Everspin 256Mb DDR3 Spin-Torque MRAM

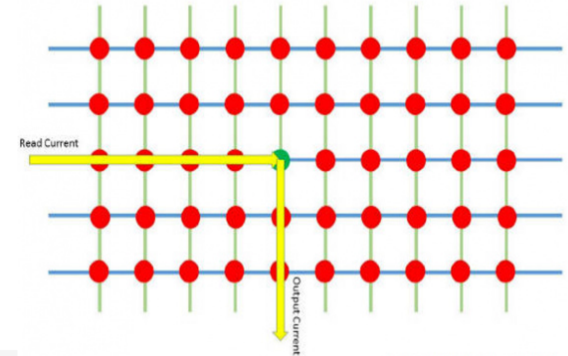
EMD3D256M[08G1/16G2]

FEATURES

- Non-volatile 256Mb (32Mb x 8, 16Mb x 16) DDR3
- Supports standard DDR3 SDRAM features
- V_{DD} = 1.5v +/- 0.075v
- Up to 667MHz ^{f}CK (1333MT/sec/pin)
- Page size of 512 bits (x8) or 1024 bits (x16)
- On-device termination
- On-Chip DLL aligns DQ, DQS, $\overline{DQS}$ transition with CK transition
- All addresses and control inputs are latched on rising edge of Clock
- Burst length of 8 with programmable Burst Chop length of 4
- Standard 10x13mm 78-Ball (x8) or 96-ball (x16) BGA Package

Desu-2017



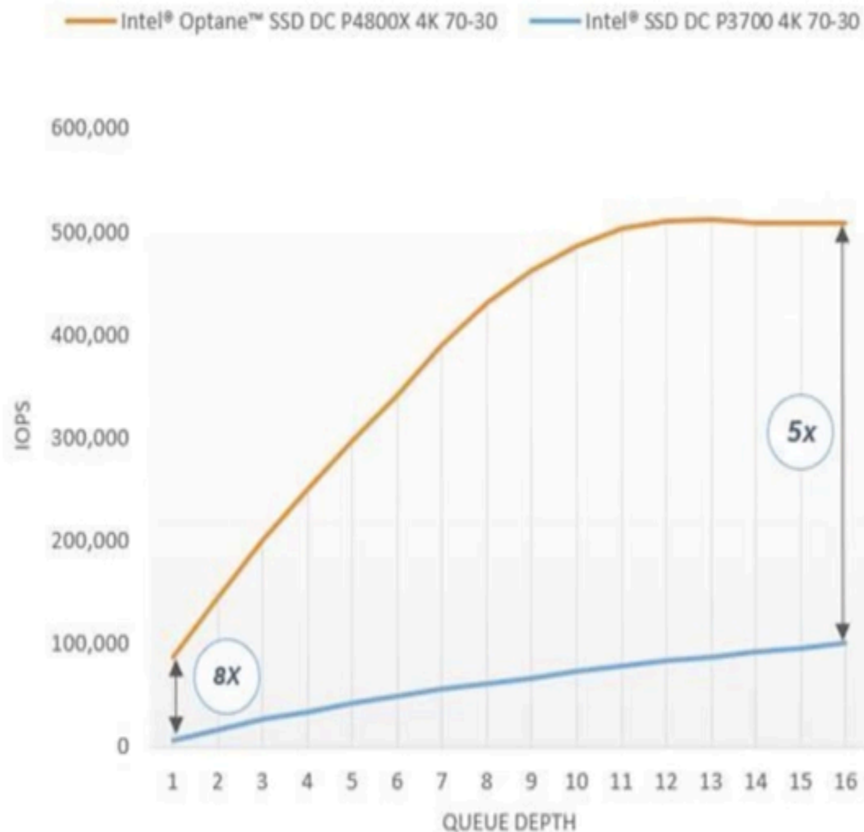


	DRAM	PCM	NAND Flash
Page size	64B	64B	4KB
Page read latency	20-50ns	~ 50ns	~ 25 µs
Page write latency	20-50ns	~ 1 µs	~ 500 µs
Write bandwidth	~ GB/s per die	50-100 MB/s per die	5-40 MB/s per die
Erase latency	N/A	N/A	~ 2 ms
Endurance	∞	$10^6 - 10^8$	$10^4 - 10^5$
Read energy	0.8 J/GB	1 J/GB	1.5 J/GB [28]
Write energy	1.2 J/GB	6 J/GB	17.5 J/GB [28]
Idle power	~ 100 mW/GB	~ 1 mW/GB	1–10 mW/GB
Density	1×	2 – 4×	4×

Desu-2017

THE MEMORY TECHNOLOGY EVOLUTION
…Is Not Enough
1947
1956
PROM
1961
SRAM
1966
DRAM
1971
EPROM
1979
Bubble Memory
1980
NOR Flash Memory
1984
NAND Flash Memory
2017
?
Desu-2017

Breakthrough Performance



+

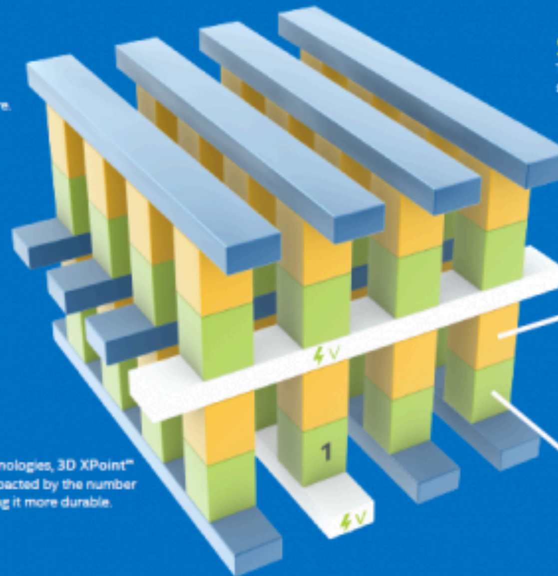
- **5-8x faster** at low Queue Depths[1]
- Vast majority of **applications generate low QD** storage workloads

Desu-2017

3D XPoint™ Technology:
An Innovative, High-Density Design

Cross Point Structure
Perpendicular wires connect submicroscopic columns. An individual memory cell can be addressed by selecting its top and bottom wire.

Stackable
These thin layers of memory can be stacked to further boost density.



Selector
Whereas DRAM requires a transistor at each memory cell—making it big and expensive—the amount of voltage sent to each 3D XPoint™ Technology selector enables its memory cell to be written to or read without requiring a transistor.

Non-Volatile
3D XPoint™ Technology is non-volatile—which means your data doesn't go away when your power goes away—making it a great choice for storage.

High Endurance
Unlike other storage memory technologies, 3D XPoint™ Technology is not significantly impacted by the number of write cycles it can endure, making it more durable.

Memory Cell
Each memory cell can store a single bit of data.

Transforming the Memory Hierarchy
For the first time, there is a fast, inexpensive and non-volatile memory technology that can serve as system memory and storage.

3D XPoint™ Technology　　Processor

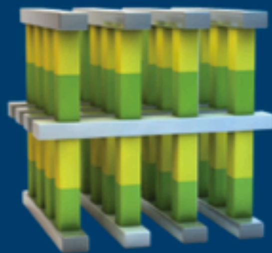
~8x to 10x Greater Density than DRAM[1]
3D XPoint™ Technology's simple, stackable, transistor-less design packs more memory into less space, which is critical to reducing cost.

DRAM　　3D XPoint™ Technology

Desu-2017

WHAT IS INTEL® OPTANE™ MEMORY?

3D XPOINT™ MEMORY MEDIA



Scalability

Cross Point Structure

High Performance

Breakthrough Material Advances

STANDARD M.2 CONNECTOR MODULAR FORM FACTOR

PCIe* Gen 3.0x2
M.2 2280
Single-sided

INTEL® RAPID STORAGE TECHNOLOGY

The two physical devices are paired into a single volume

Files needed for important tasks are immediately recognized and accelerated

Over time, frequently used files and applica-tions are monitored and accelerated as well

INTEL® OPTANE™ MEMORY

M.2 2280

16GB 32GB

FOR 7TH GEN INTEL® CORE™ PLATFORM

Desu-2017