TDK

Basic Principles, Challenges and Opportunities of STT-MRAM for Embedded Memory Applications

Luc Thomas

TDK- Headway Technologies,
463 S. Milpitas Boulevard, Milpitas CA 95035, USA

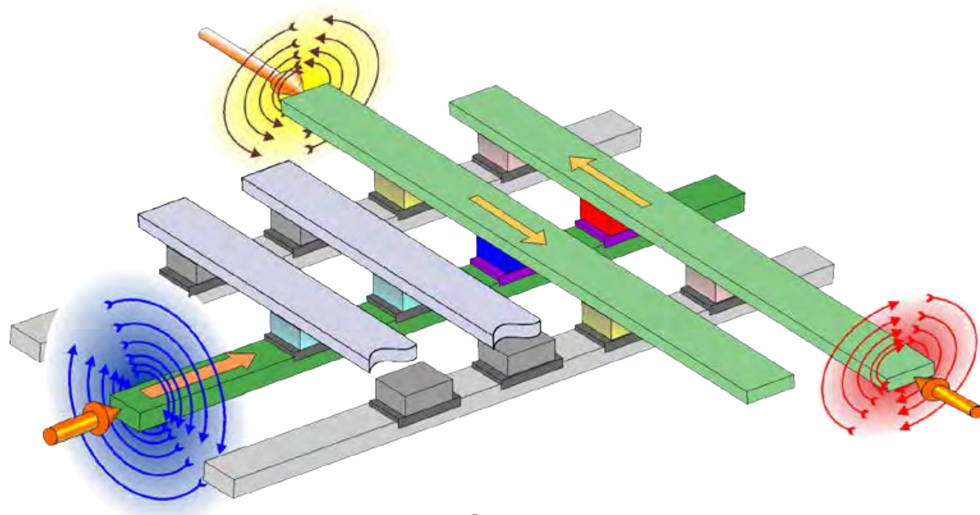
MRAM Team at TDK - Headway Tech.

Guenole Jan, Son Le, Santiago Serrano-Guisan
Yuan-Jen Lee, Huanlong Liu, Jian Zhu,
Jodi Iwata-Harms, Ru-Ying Tong, Sahil Patel,
Vignesh Sundar, Dongna Shen, Yi Yang, Renren He,
Jesmin Haq, Jeffrey Teng, Vinh Lam, Paul Liu,
Yu-Jen Wang, Tom Zhong, and Po-Kang Wang.

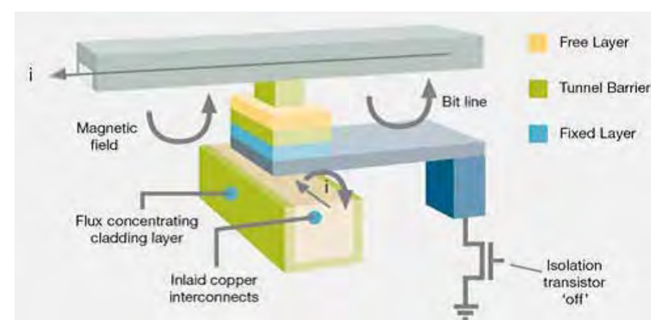
Magnetic Random Access Memories

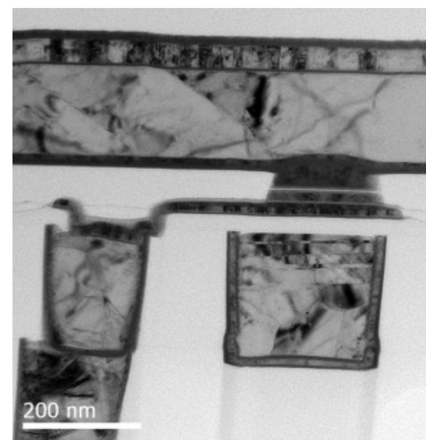
➔ More than 20 years ago: Field-MRAM

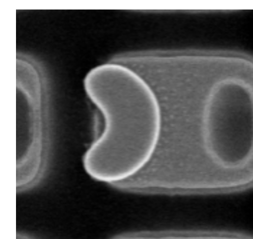
1st research program: IBM / Motorola (1995)
1st product: Freescale / Everspin (2006)



From S. Parkin and K. Roche IBM

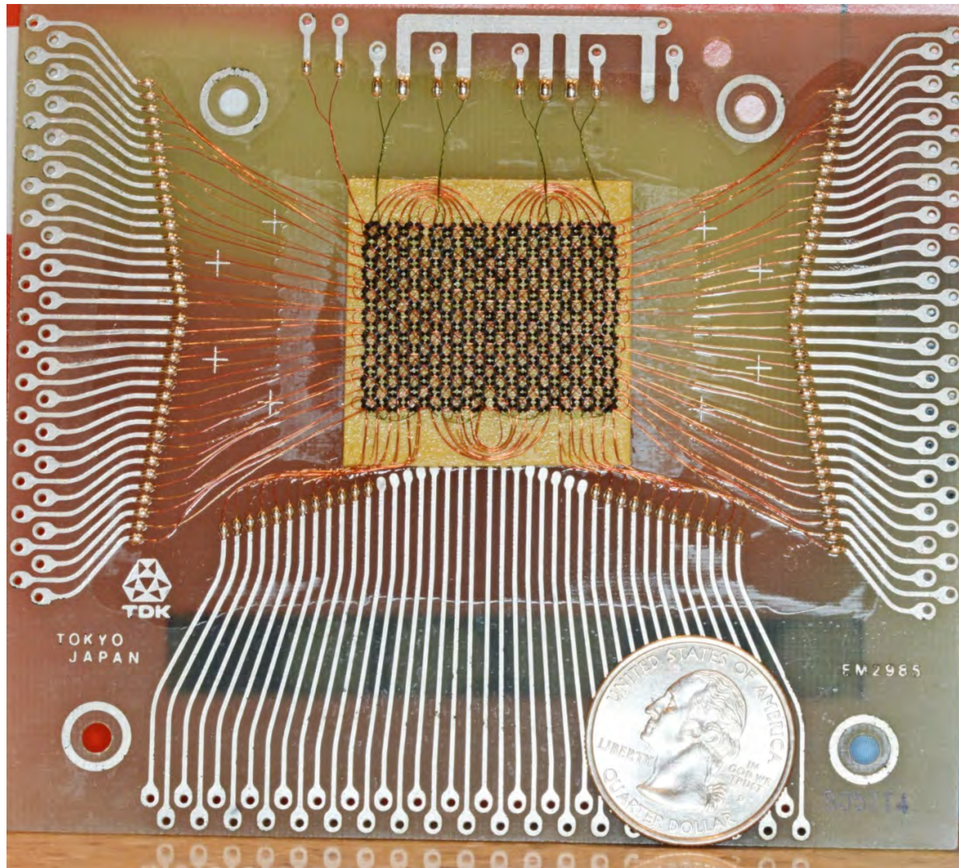




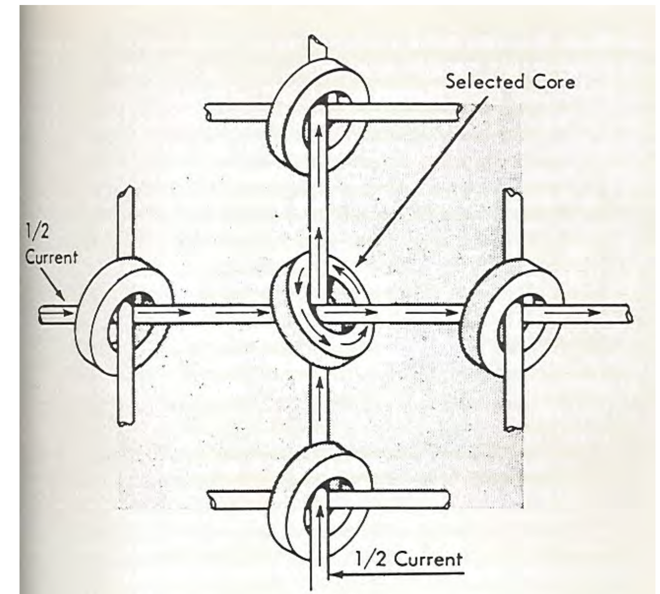


60 years ago: TDK first foray in MRAM technology

➔ TDK's 18x24 bit Magnetic Core Memory



Source: wikipedia.org/wiki/Magnetic-core_memory



Source: columbia.edu/cu/computinghistory/core.html

➔ MRAM was the predominant computer memory from the 50's to the 70's

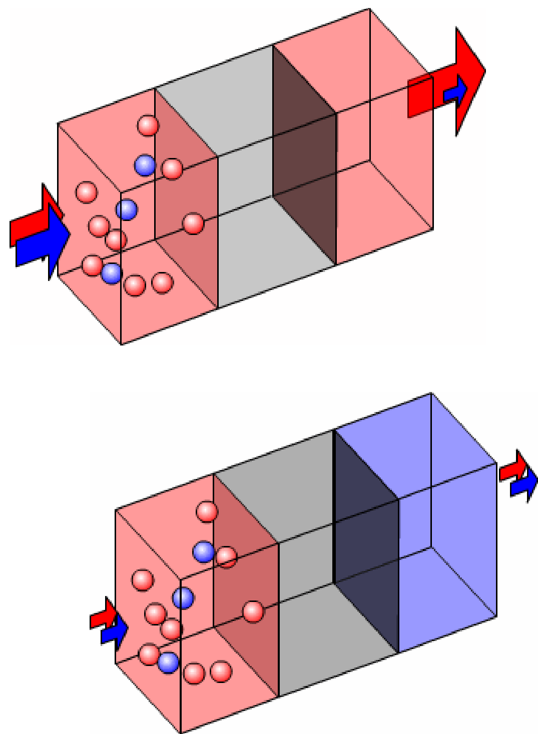
Outline

- ➔ **Basic principles of STT-MRAM**
- ➔ STT-MRAM integration
- ➔ STT-MRAM in emerging memory landscape

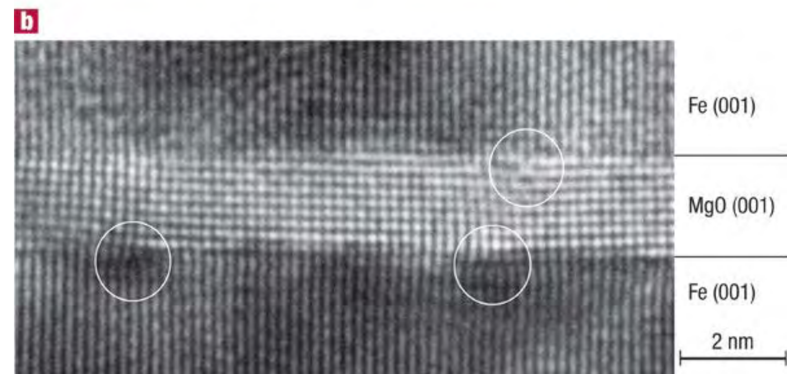
Magnetic Tunnel Junction (MTJ) device

➔ Two ferromagnetic electrodes separated by a thin MgO tunnel barrier

➔ Tunnel Magnetoresistance (TMR): device resistance depends on the relative orientation of the magnetization of the two magnetic electrodes



From S. Parkin and K. Roche IBM

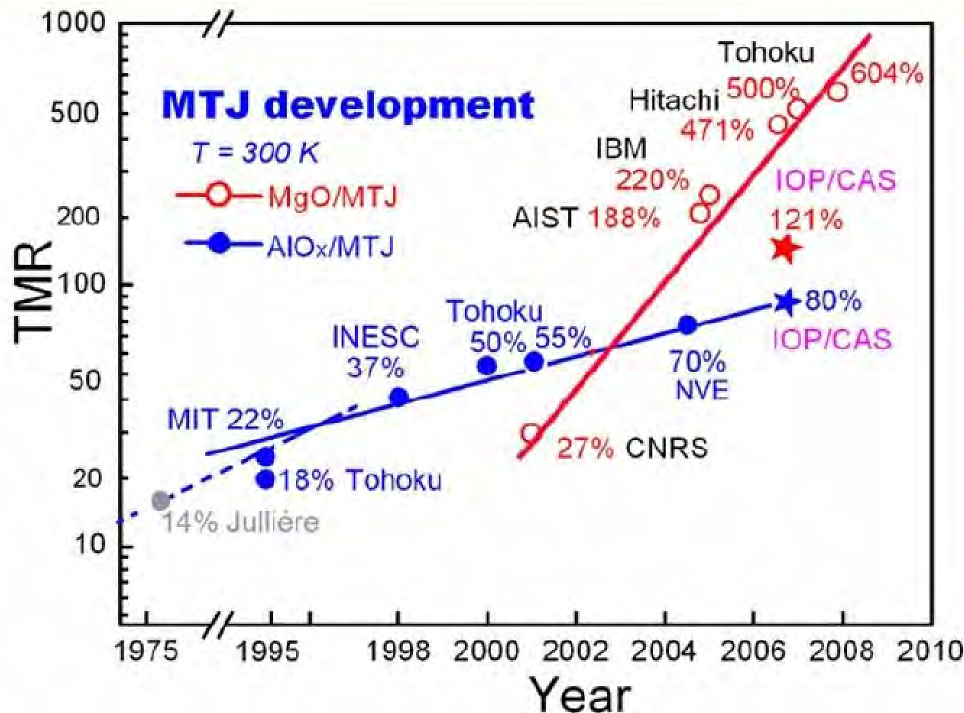


Yuasa et al. (AIST) Nature Materials2004

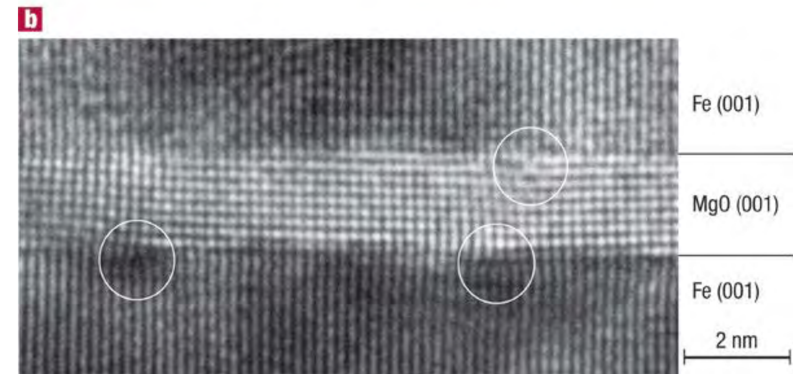
Magnetic Tunnel Junction (MTJ) device

➔ Two ferromagnetic electrodes separated by a thin MgO tunnel barrier

➔ Tunnel Magnetoresistance (TMR): device resistance depends on the relative orientation of the magnetization of the two magnetic electrodes



Reproduced from website of MultiDimension Technology Co.,Ltd.

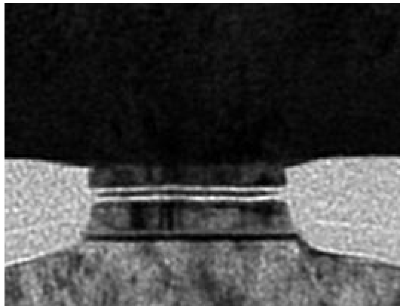


Yuasa et al. (AIST) Nature Materials2004

Perpendicular Magnetic Anisotropy (PMA) MTJ

- ➔ PMA is needed f … ciency
- ➔ PMA is based or … d CoFeB

  Ikeda et al., Nature
- ➔ Free layer sandv … e free layer for enhanced anisot
- ➔ Dual reference la … nced stability



Ru(5)
Ta(1)
Recording layer
MgO
CoFeB (1.0)
Ta(0.45)
CoFeB (1.6)
MgO
Reference layer
CoFeB(1.0)
Ta(0.3)
$[Co(0.4)/Pt(0.4)]_2$
Co(0.4)
Ru(0.42)
Co(0.4)
$[Co(0.4)/Pt(0.4)]_6$
Pt(5)
Ta(5)
Si/SiO_2 sub.

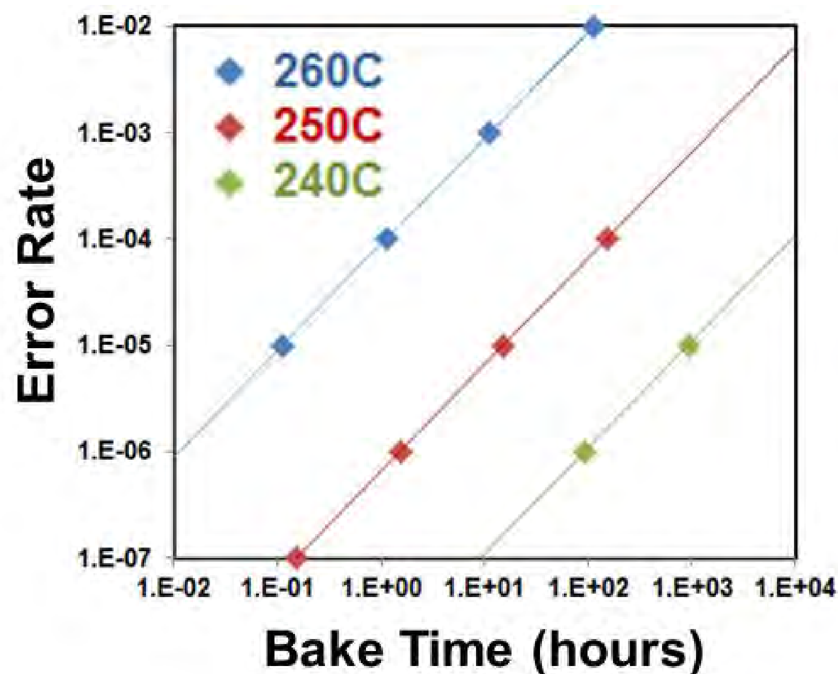
Ikeda et al., IEDM2014

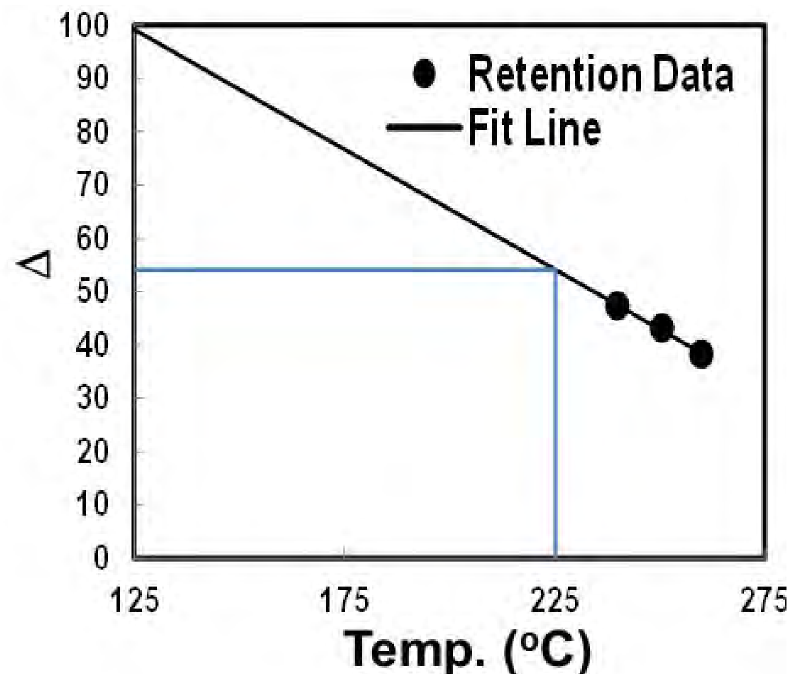
Free Layer
Pinned Layer 1
Pinned Layer 2

High data retention in PMA-MTJs

➔ Developed a MTJ stack of high PMA and thermal stability to satisfy solder reflow requirement of 260ºC for 90 seconds (2016 VLSI TSMC/TDK)

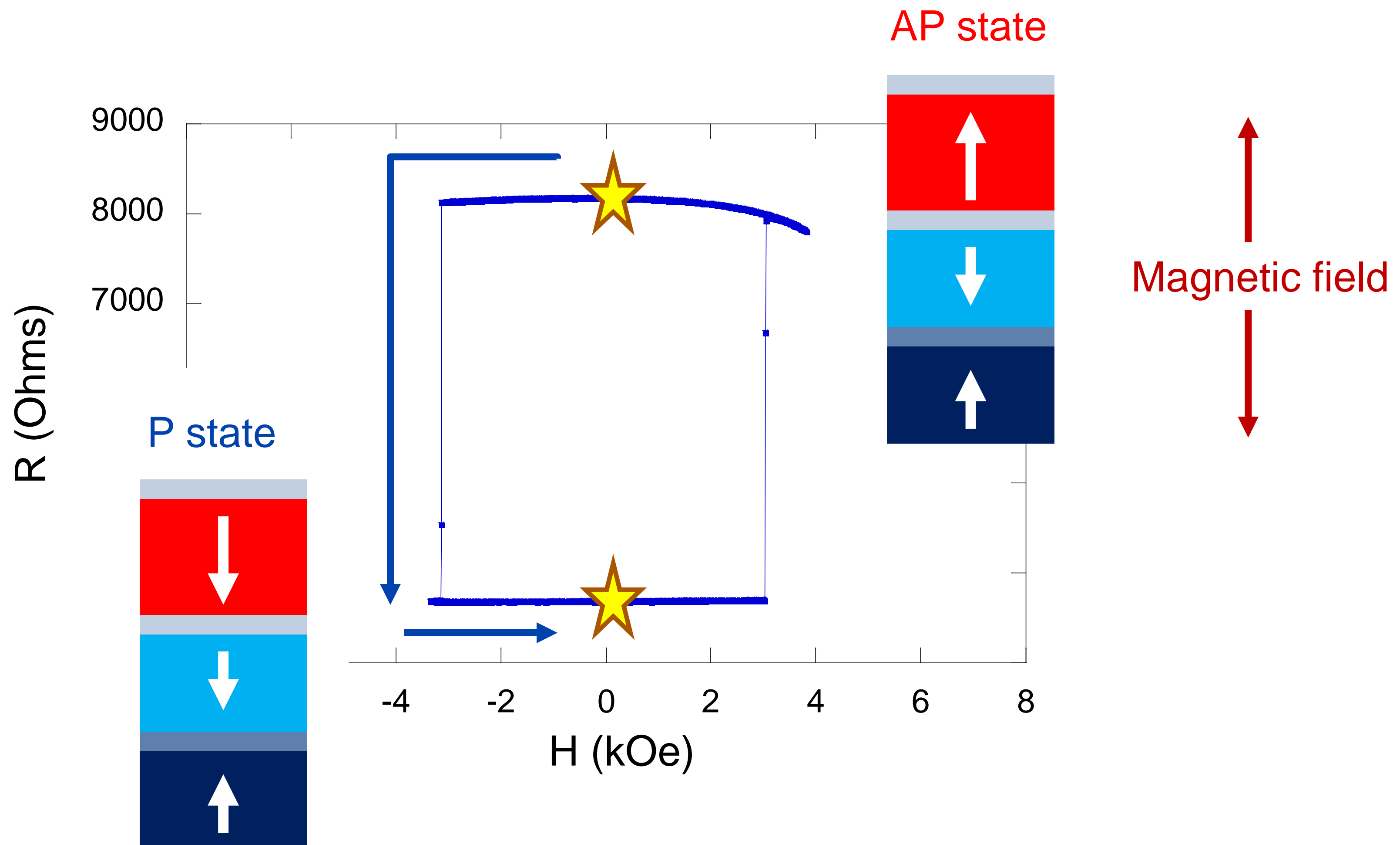
➔ Method of projecting error rate from chip level data in ppm regime





1ppm 10 years retention at 225ºC

Resistance vs magnetic field hysteresis loops

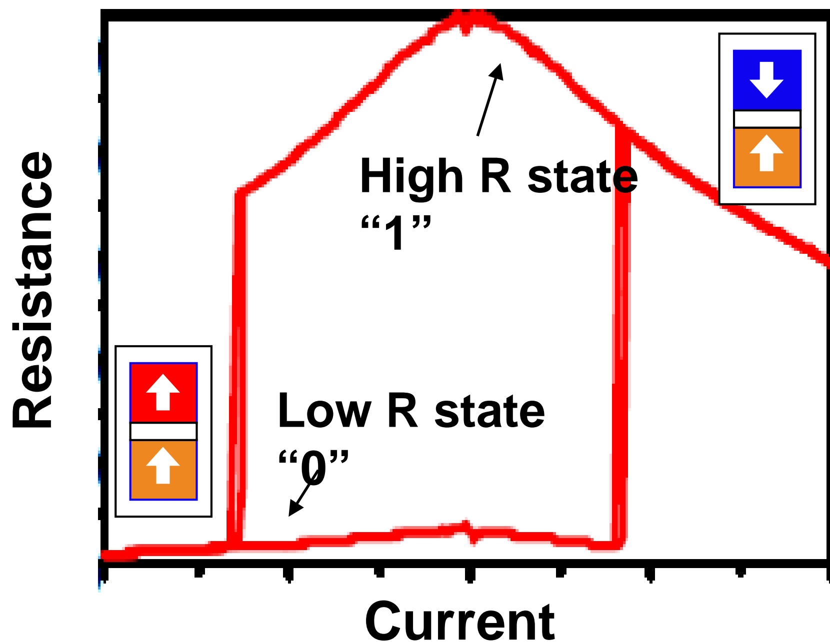


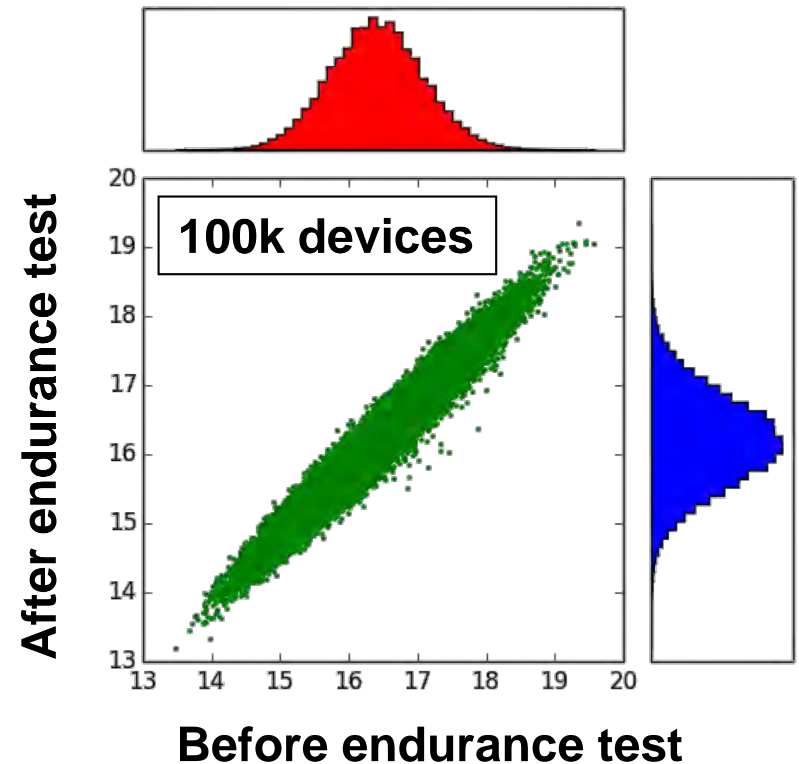
➔ Two well-defined resistance states depending on orientation of magnetic electrodes

Reading with Tunnel Magnetoresistance

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

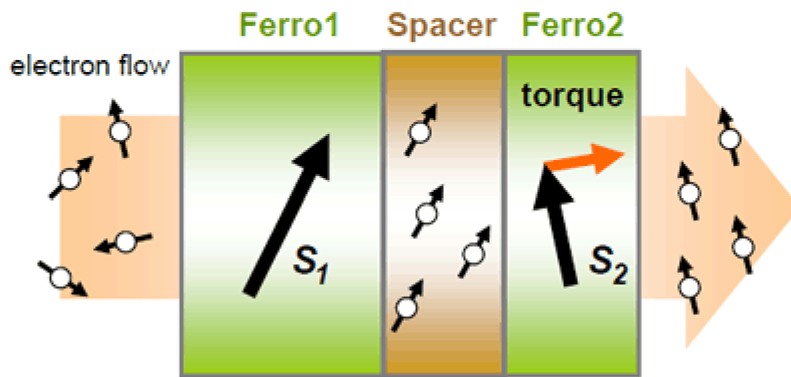
➔True Binary device: no resistance drift of the 2 resistance state even after repeated cycling at maximum drive current





Writing with Spin-Transfer Torque

➔ Transfer of spin-angular momentum from polarized conduction electrons to electrodes magnetization

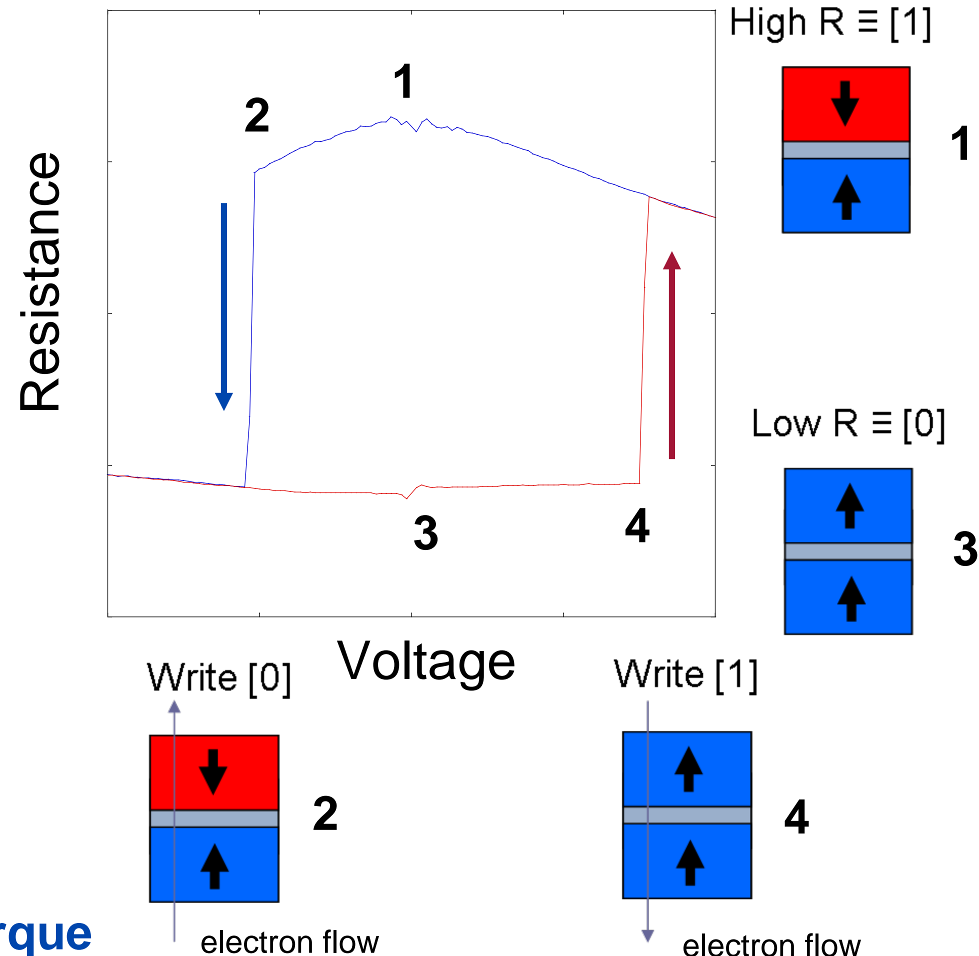


Reproduced from Quantumwise.com

Phenomenon discovered in 1996
by two theoreticians:
John Slonczewski (IBM)
Luc Berger (Carnegie Mellon)

Read:

Tunnel Magnetoresistance



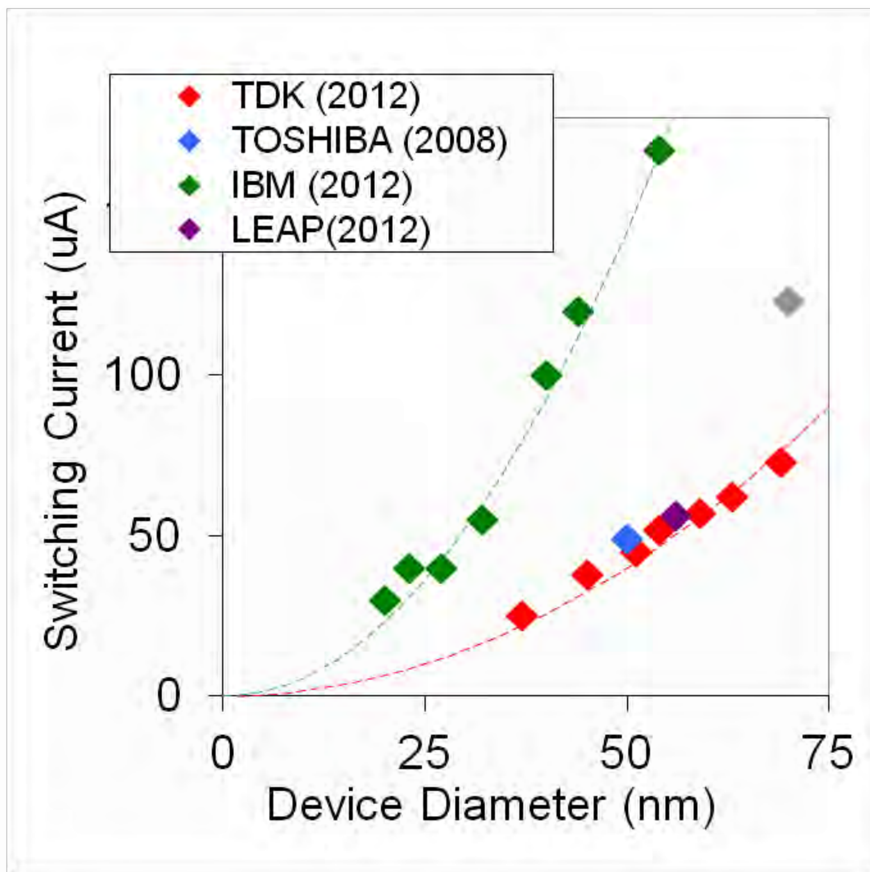
Write:

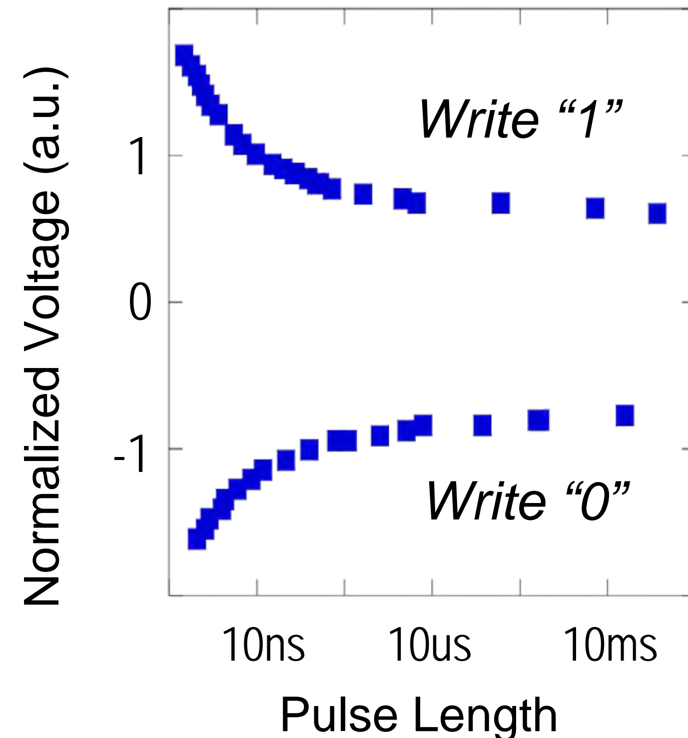
Spin Transfer Torque

Trade-offs of STT writing

➔ Switching Current scales with area (constant current density)
- smaller device -> smaller current requirement

➔ Current inversely proportional to pulse width
- faster -> higher current requirement





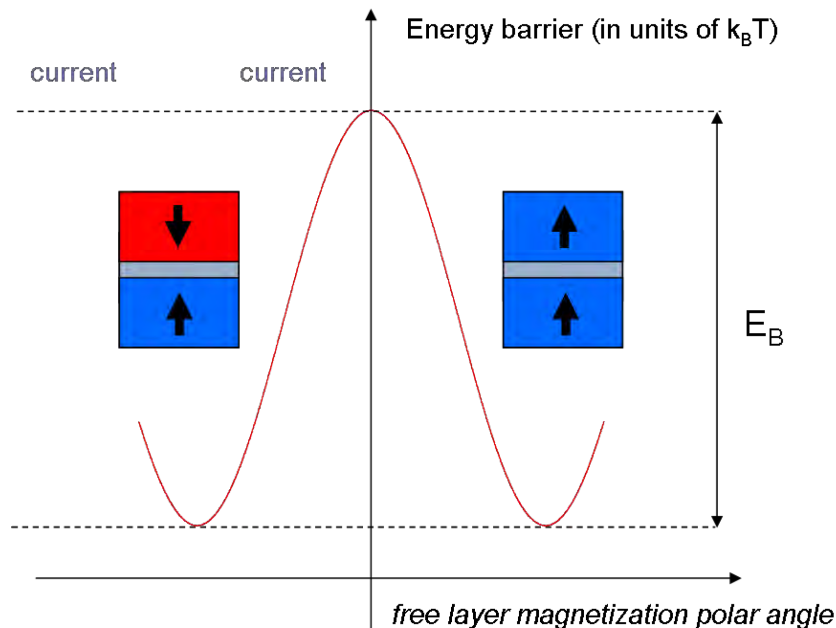
Trade-offs of STT writing (cont'd)

➔ Write current scales with energy barrier for data retention

Energy barrier: $E_B \sim K_u V$

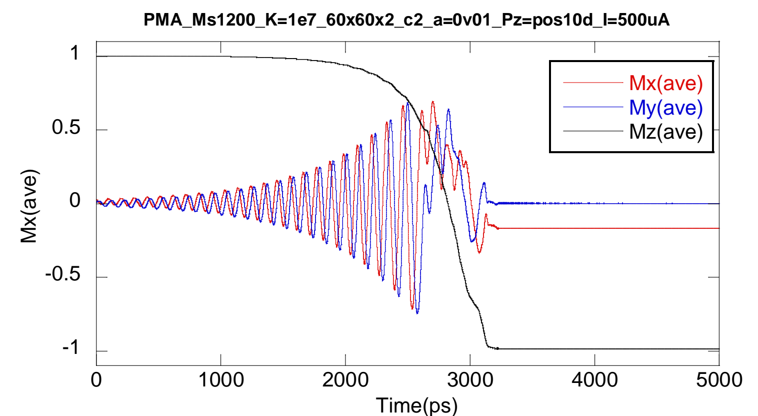
Write current: $I_{c0} = (4e/\hbar)\,(\alpha/P)\,E_B$

STT efficiency: E_B/I_{c0} ~ 1-2 in $k_BT/\mu A$



➔ Writing is probabilistic

- STT vanishes for parallel alignment of PL and FL
- Switching time inversely proportional to angle between PL and FL
- Thermal fluctuations provide initial 'kick'

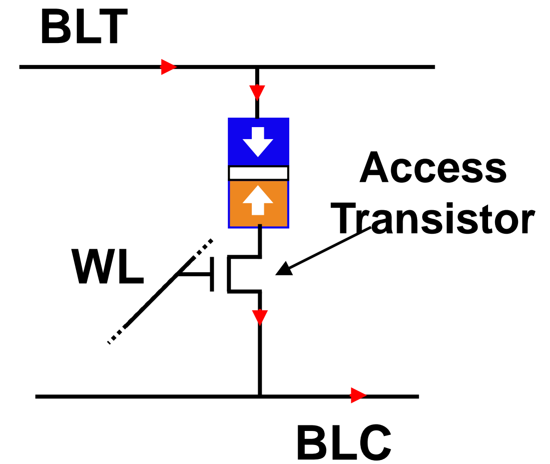


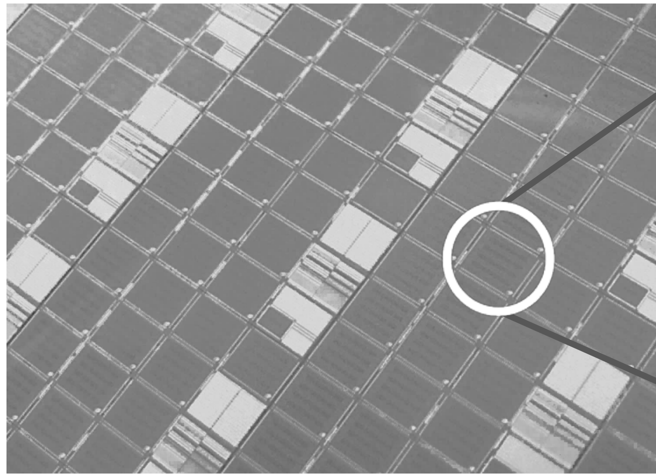
Outline

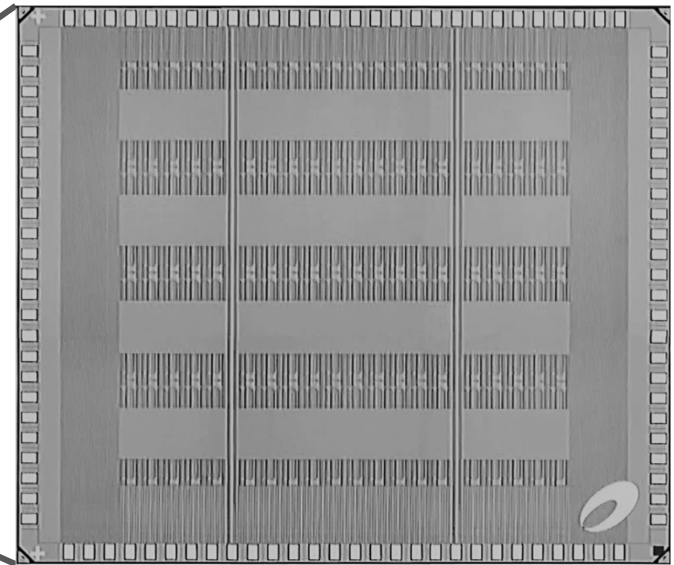
- ➔ Basic principles of STT-MRAM
- ➔ **STT-MRAM integration**
- ➔ STT-MRAM in emerging memory landscape

Integration of 8 Mb test chips at TDK - Headway

- 8Mbits (16x512k) 1T-1MTJ
- IBM's 90nm CMOS technology
- $50F^2$ cell size
- Sense Amplifiers for reading
- Redundancy and 2bit ECC
- FEOL in IBM foundry
- BEOL in TDK-Headway's fab

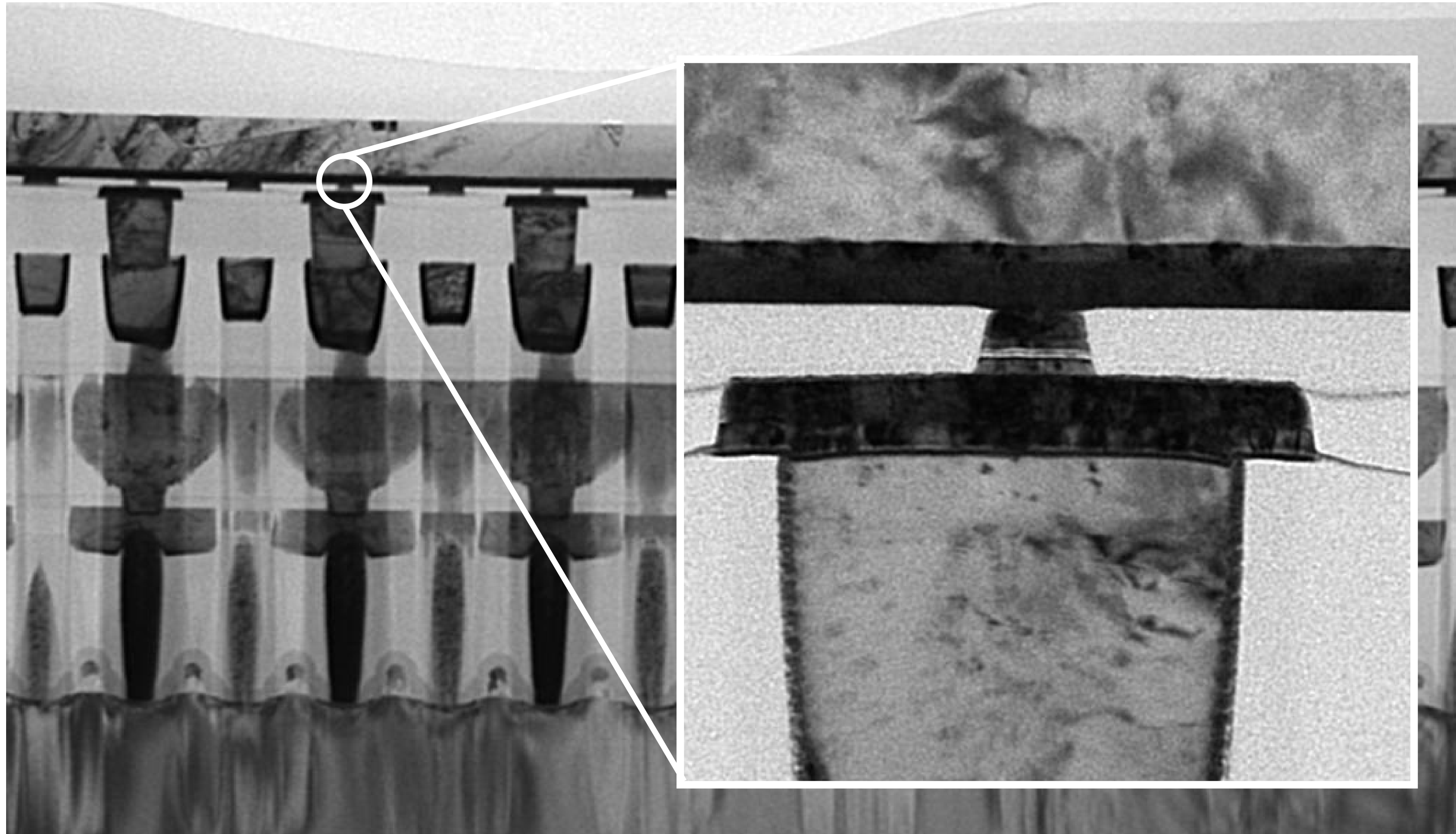






STT MRAM process integration

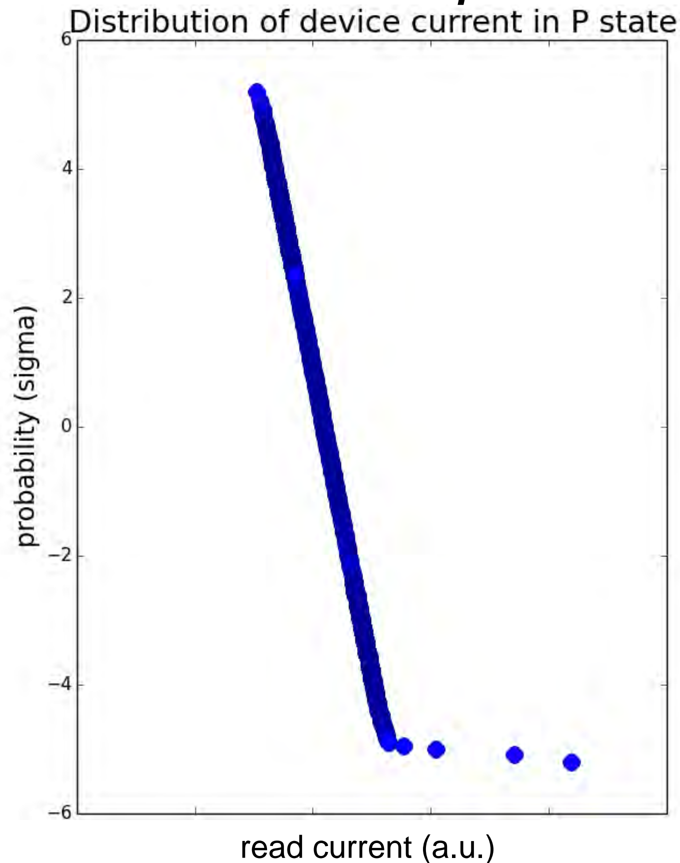
- MRAM only add three additional layers (MTJ and electrodes) to standard CMOS BEOL: 3 to 4 mask adder
- MTJ stack is about 20 nm thick, can be easily integrated into CMOS backend process



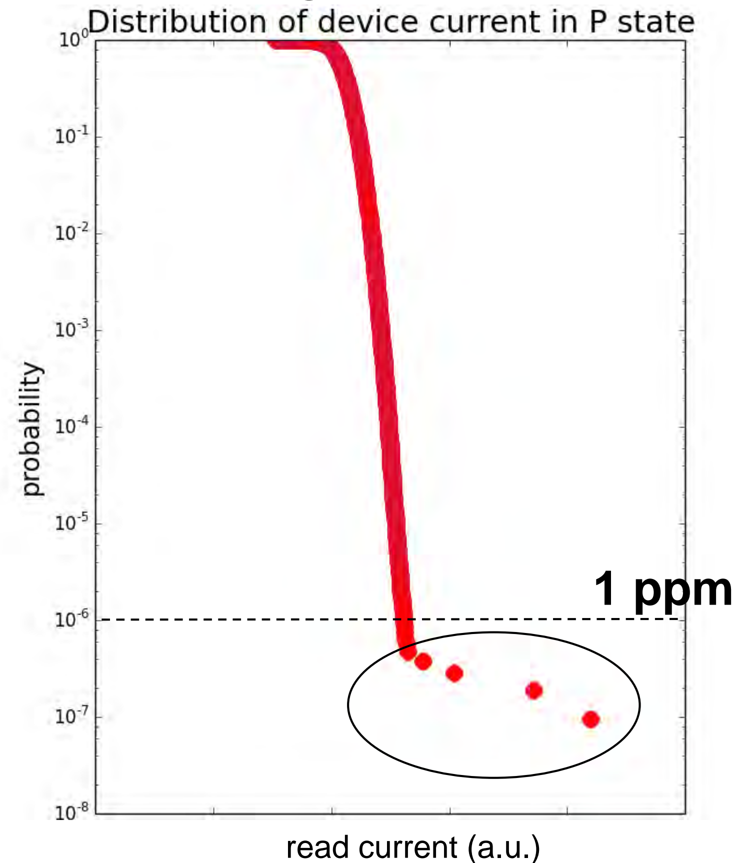
Defect rate of 8 Mb chip

- Distribution of device current in the P state

Quantile plot



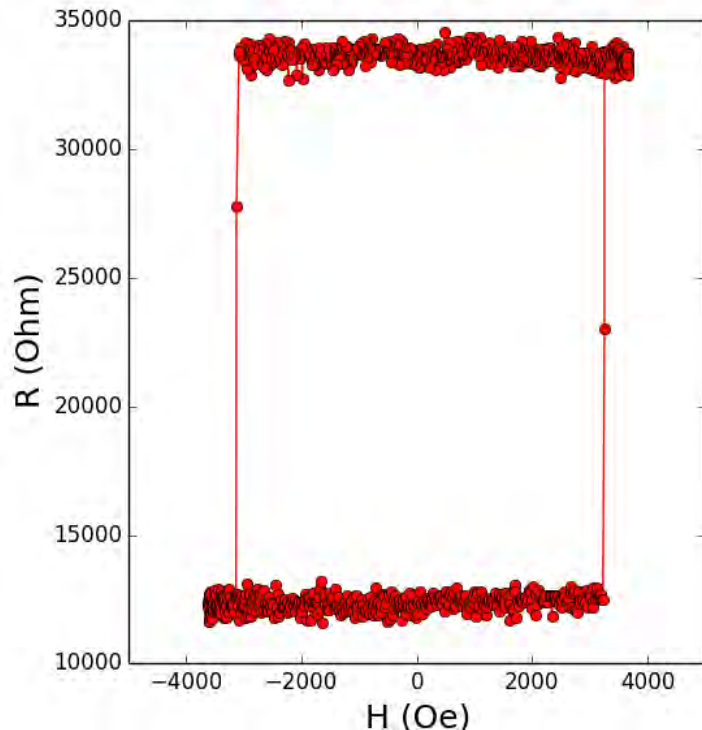
Log scale



➔ less than 0.4 ppm defect rate

400C annealing after MTJ patterning

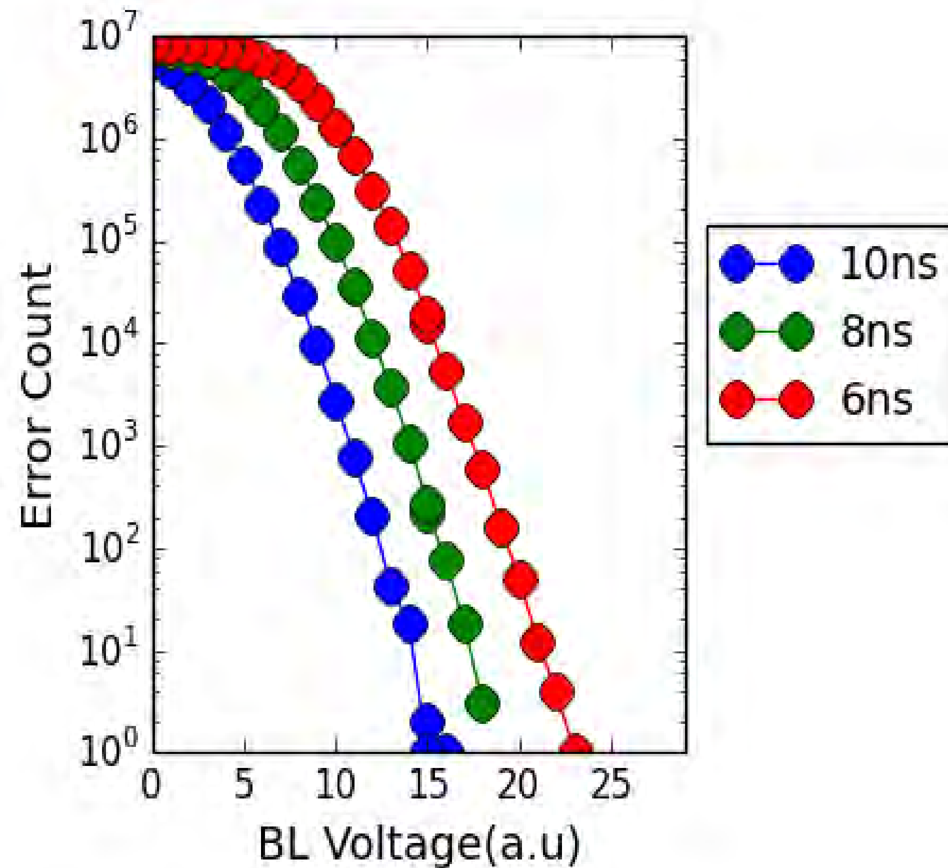
- ➔ 400C BEOL process can add up to several hours, depending on how many metal layers on top of MTJ
- ➔ Elemental movements and morphology changes can degrade anisotropy, exchange coupling, and defect level
  - - selection of materials, diffusion barrier and interface/growth quality
  - - Thorough engineering needed for electrodes, film stack, process, encapsulation

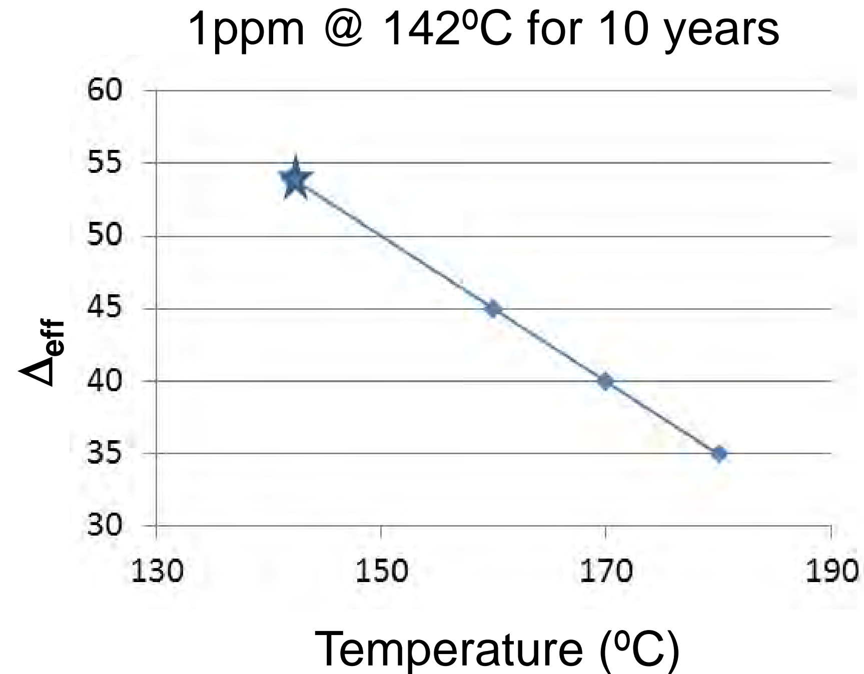


- ➔ **2.5 hours @400ºC after MTJ etching**
- ➔ **Diameter ~ 30 nm (electrical)**
- ➔ **DRR = 175%**
- ➔ **RA of 8.5 Ω-µm²**
- ➔ **H_C = 3300 Oe with no offset**

Error free writing in chip level (TDK VLSI2014 & 2016)

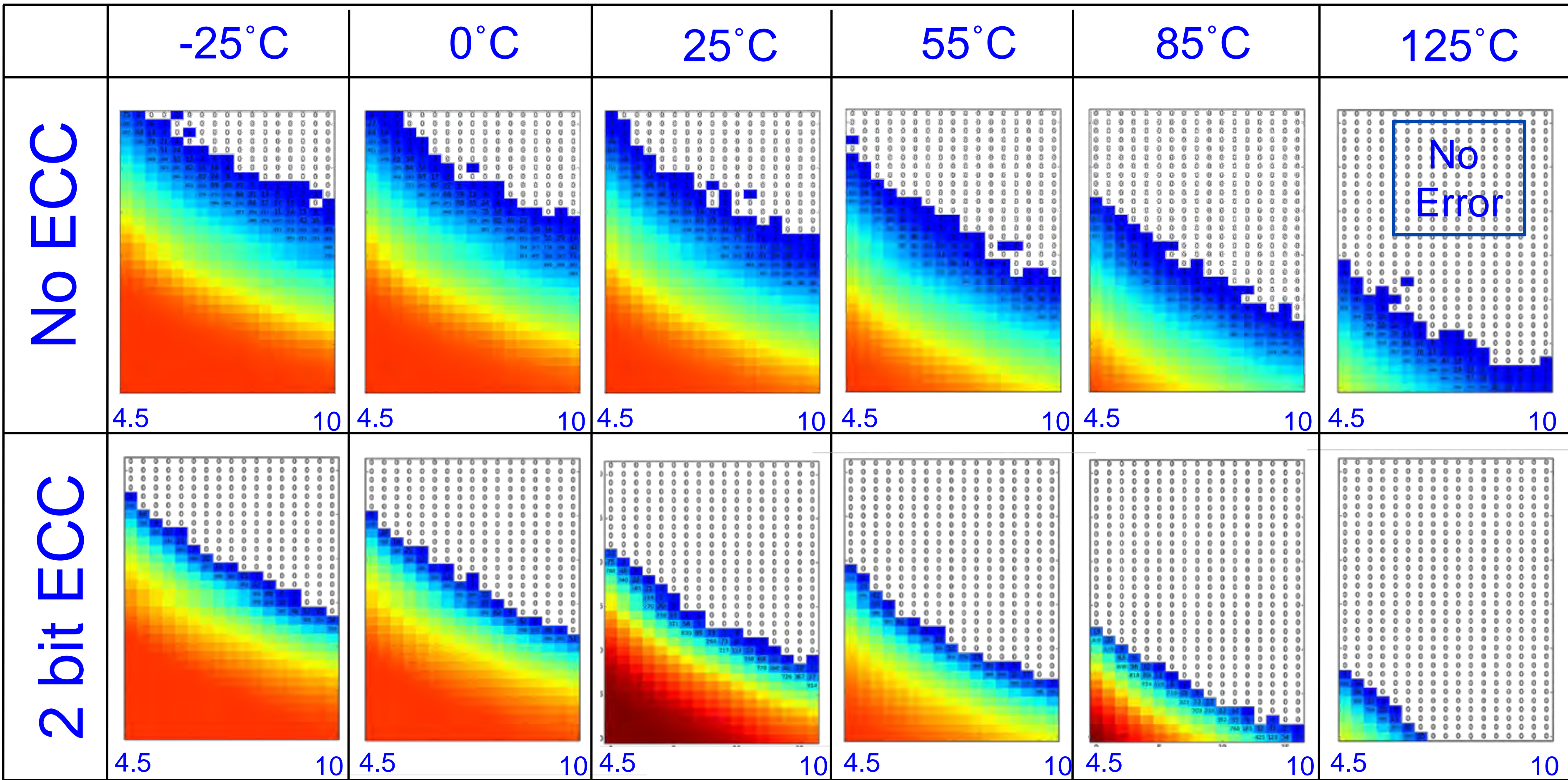
➔ Error free writing on 8 Mb chips without ECC
- Down to 6 ns write pulse
- While keep data retention to 142ºC for 10 years





Temperature dependence (TDK VLSI2014)

Fast operation down to 4.5 ns demonstrated over wide temperature range



Outline

- Basic principles of STT-MRAM
- STT-MRAM integration
- **STT-MRAM in emerging memory landscape**

STT-MRAM vs other memory technologies

Stand Alone Memory Position in 2014 : commercial products performances

	Emerging Memory			Established Memory		
	STT MRAM	PCM	RRAM	SRAM	DRAM	Flash NAND
Non-Volatile	YES	YES	YES	NO	NO	YES
Endurance (Nb cycles)	High (10^{12})	Medium (10^{8})	Low (10^{6})	High (10^{15})	High (10^{15})	Low (10^{5})
2014 latest technological node produced (nm)	90 nm	45 nm	130 nm	10 nm	30 nm	15 nm
Cell size (cell size in F^2)	Medium (6-12)	Medium (6-12)	Medium (6-12)	Very large (150)	Small (6-10)	Very small (4)
Write speed (ns)	High (10 ns)	Medium (75 ns)	Very high (10 ns)	High (5-10 ns)	High (10 ns)	Low (10,000 ns)
Power consumption	Medium/low	Medium	Low (3- 5 pJ/bit)	Very low	Low	Very high
2014 price ($/Gb)	High ($100 - 50/Gb)	Medium (few $/Gb)	High ($5,000 /Gb)	Low ($1/Gb)	Low ($1/Gb)	Very Low ($0.05/Gb)
Suppliers	Everspin	Micron, Samsung	Adesto	Qualcomm, Intel	Samsung, Micron, SK Hynix	Samsung, Micron, Toshiba, SK Hynix

- Emerging memory has distinctive technical features (endurance, speed, non-volatility), but price and scalability are obstacles to competing with dominant DRAM, NAND or SRAM memory.

© 2013- 10 YOLE Développement

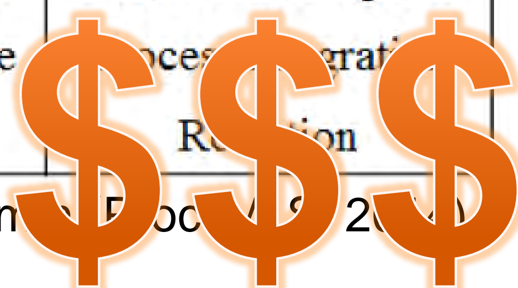
STT-MRAM requirements

➔ Critical requirements depend on application

Priority	Embedded STT-MRAM		Standalone STT-MRAM
	Unified NVM	NV Working RAM	NV DRAM NV SRAM
Higher ↑	Retention Cell Size/Chip Size Density Process Integration Endurance Performance/Speed	Performance/Speed Endurance Density Process Integration Cell Size/Chip Size Retention	Cell Size/Chip Size Density Endurance Performance/Speed Process Integration Retention

Cost!

from S.H Kang, Qualcom[illegible]



STT-MRAM Challenge

→ 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 (28,20,14/16nm,…)

  MTJ < 20 nm

  Write current < 20 μA

  TMR ~ 300%

  Kent & Worledge, Nature Nano (2015)

- Embedded Flash / DRAM : cell size **~40-50F²**

  MTJ ~ 40-100 nm

  Write current > 100 μA

  TMR > 100%

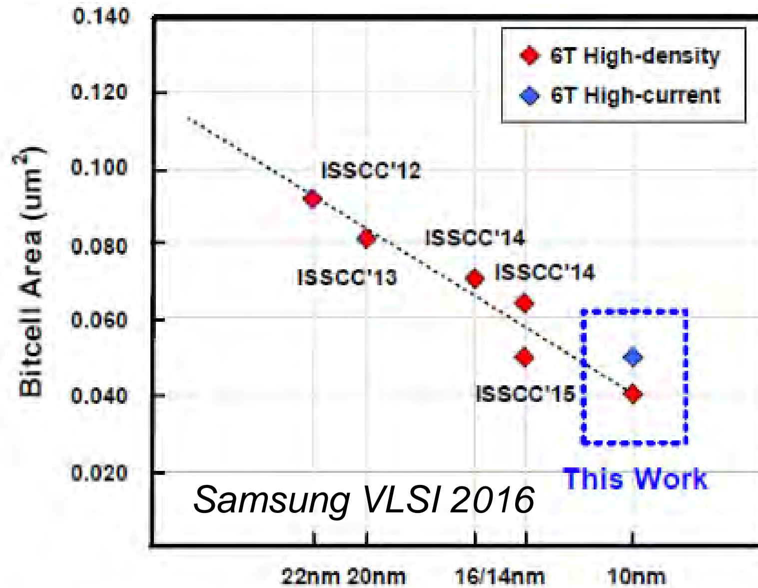
Embedded STT-MRAM is cheaper and better!

- Lower cost
  - Similar or Smaller bit cell size
  - Very few added mask layers
  - Does not interfere with CMOS transistor performances (as a add-on in the backend metal layers)

	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

- Almost "universal memory"
  - Combines non-volatility, high speed, and infinite endurance
  - Can replace eFlash, eDRAM, and last-level cache (LLC) SRAM
  - Efficient system architectures, without moving data between code storage, and working memory, and data storage
- Higher energy efficiency (longer battery life)
  - mobile and IoT applications have low duty cycles and need fast wake-up and low standby power

6-Transistor SRAM scaling challenge



[2]: ISSCC 2014 T.J. Song

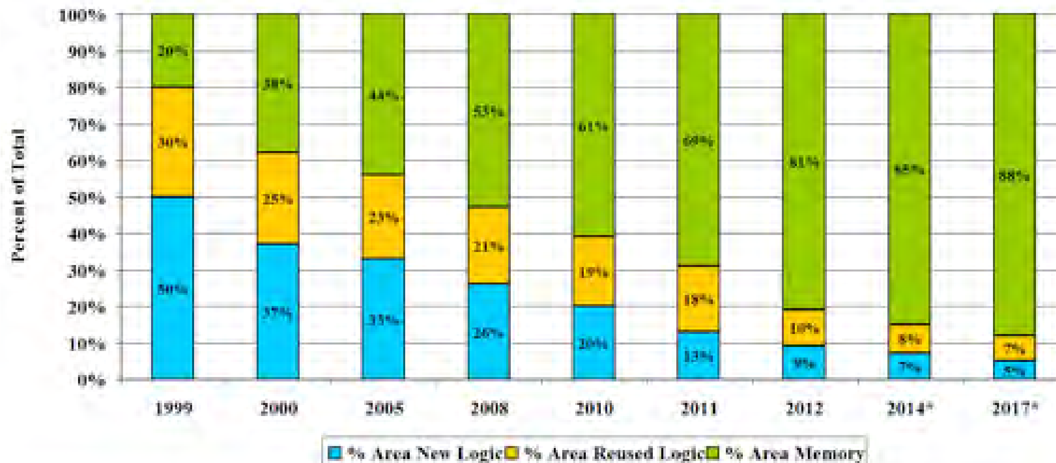
Technology Nodes

- 22nm to 10 nm node:

- Expected area scaling: 4.8X
- Actual scaling: ~ 2X

- 400F^2 at 10nm vs 52F^2 at 40nm

- Complex design limits scaling



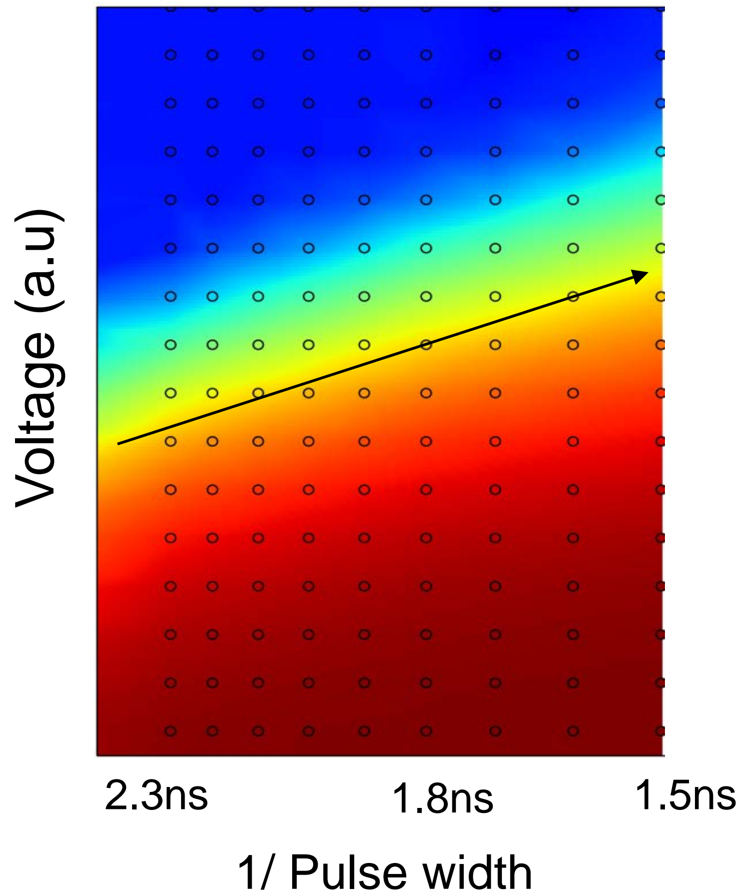
Advanced Performance Multicore SoC

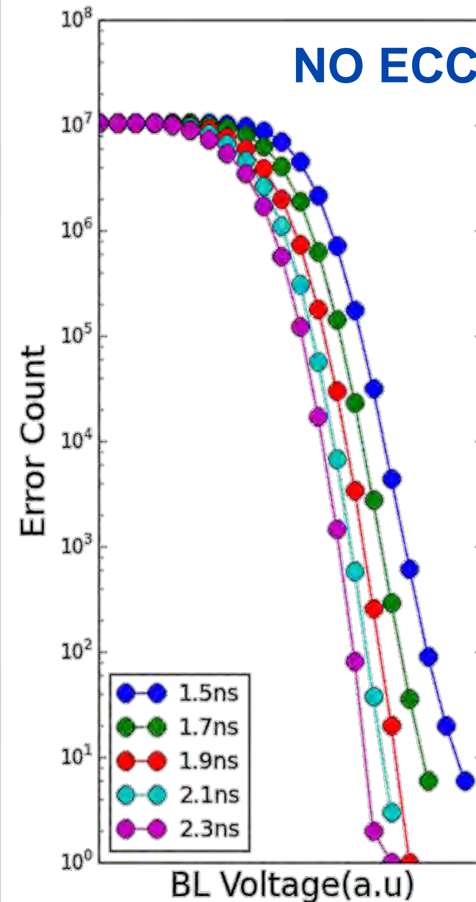
Source: Semico Research Corp.

Dramatic increase of the area occupied by memory vs logic in performance SoC and CPU's

Opportunity for eMRAM as Last Level Cache

- ➔ Compact design 1T-1MTJ
- ➔ 8 Mb written without error with 1.5 ns write pulse





TDK
VLSI 2016

Summary

➔ STT-MRAM combine low write current, data retention and write speed, and is compatible with BEOL processes.

➔ Working chips have been demonstrated

➔ MTJ device can be tailored to specific applications that require data retention or speed,

➔ Great opportunity for embedded applications from eFlash to SRAM replacement (both Samsung and TSMC have announced production)

➔ Many challenges remain: writing efficiency, read margin (TMR), process control (tight pitch, uniformity), …

1970: Magnetic memories lose the war to Silicon

Circa 1970 – Intel corporation - Computer history museum