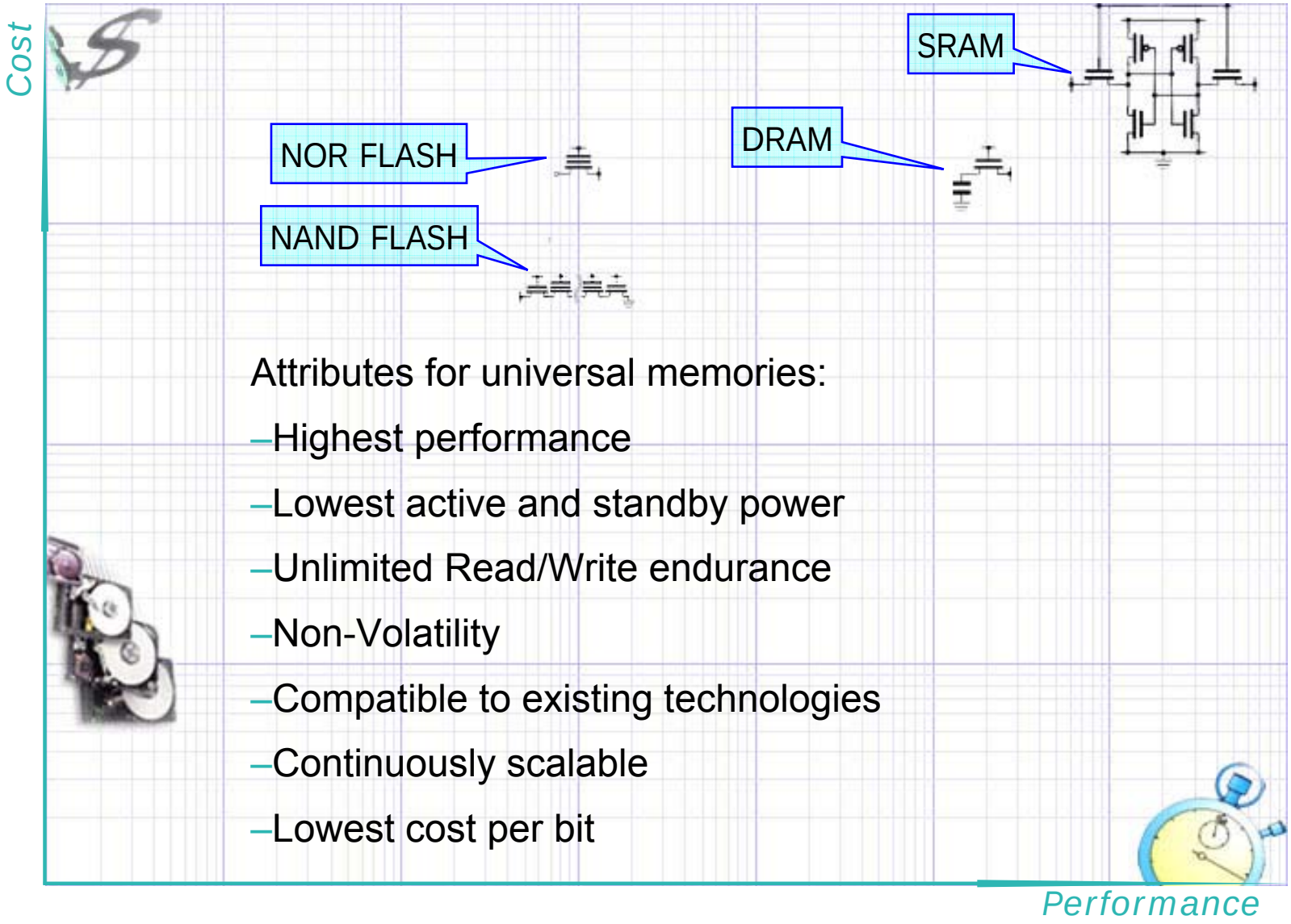
IBM Research

Prospects for Solid State Data Storage: Beyond Flash Memory and the Hard Disk Drive

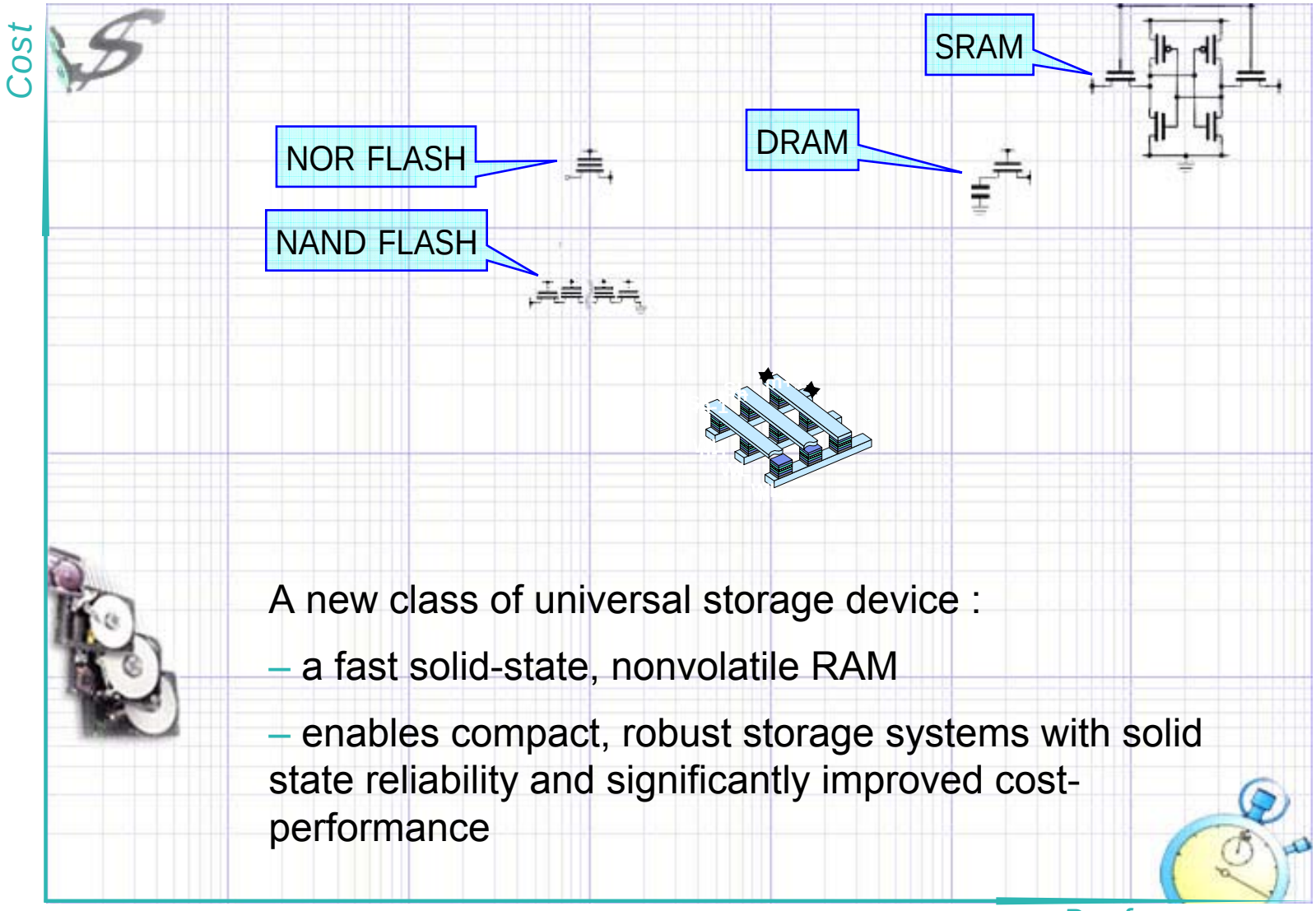
Gian-Luca Bona

gianni@us.ibm.com
IBM Research,
Almaden Research Center

Incumbent Semiconductor Memories



Incumbent Semiconductor Memories

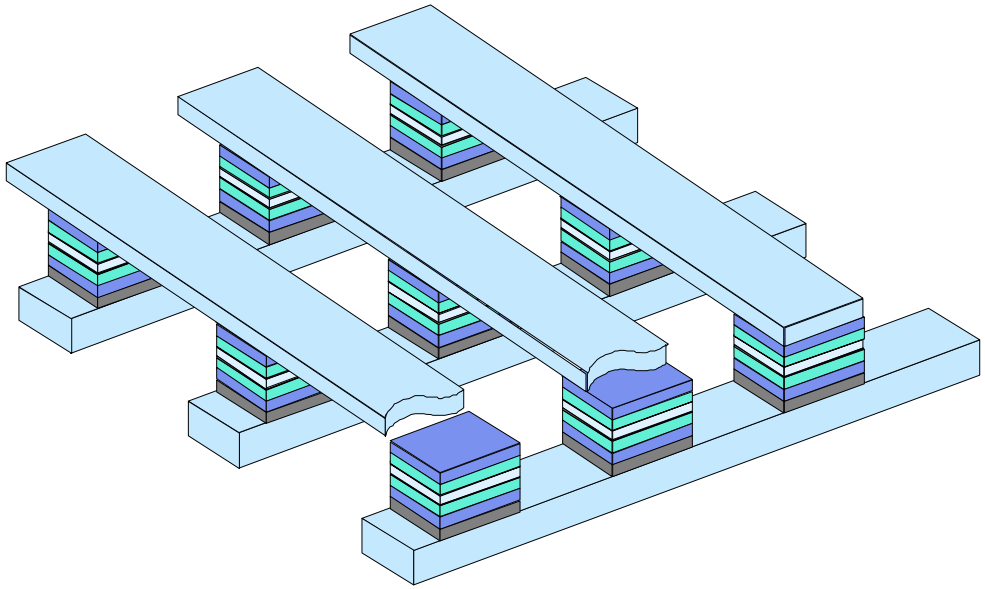


A new class of universal storage device :

– a fast solid-state, nonvolatile RAM

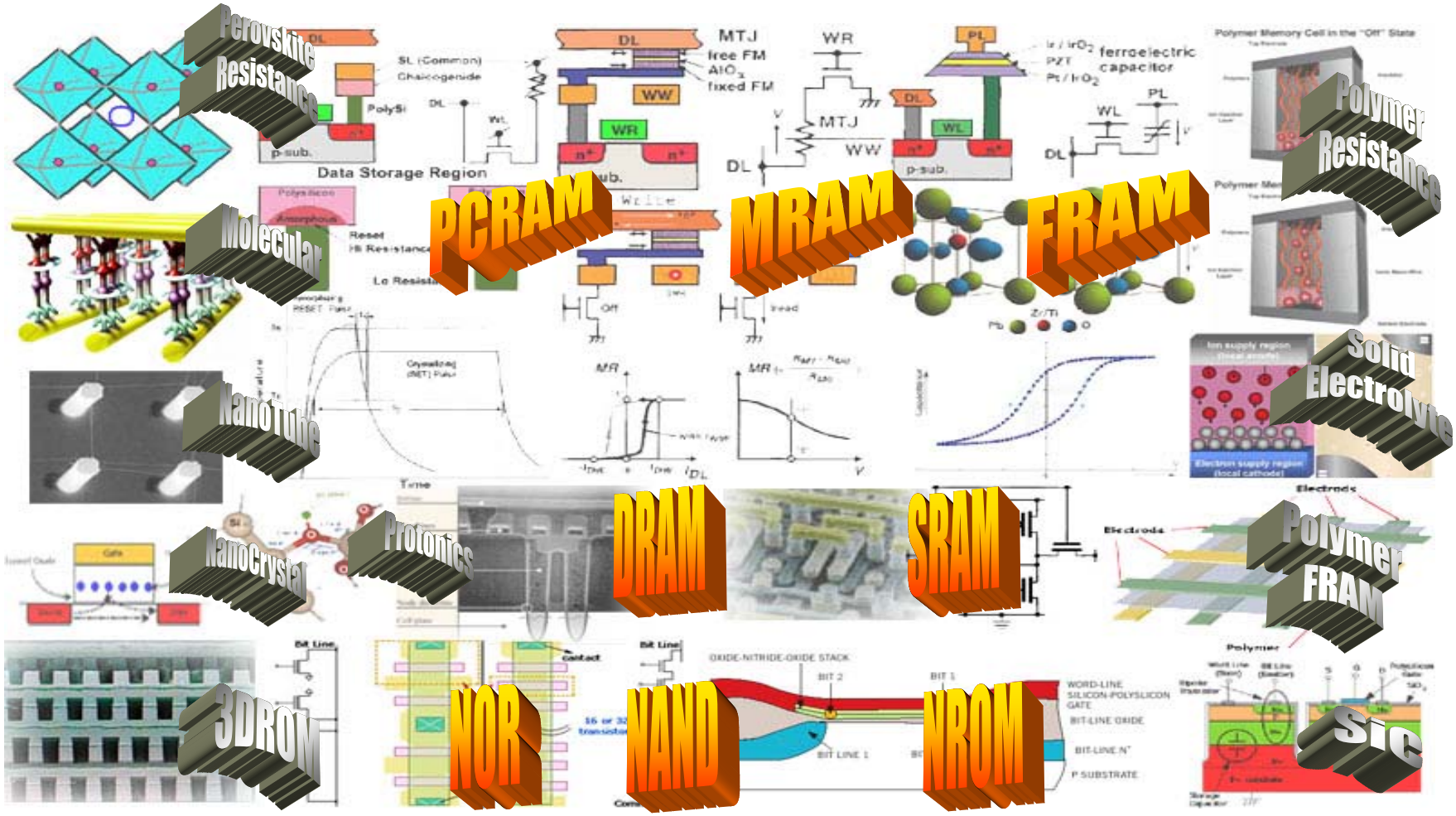
– enables compact, robust storage systems with solid state reliability and significantly improved cost-performance

Non-volatile, universal semiconductor memory



- Everyone is looking for a dense (cheap) crosspoint memory.
- It is relatively easy to identify materials that show bistable hysteretic behavior (easily distinguishable, stable on/off states).

The Memory Landscape



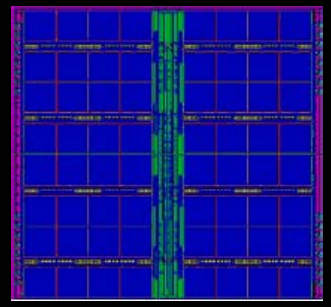
Emerging Memory Technologies

Memory technology remains an active focus area for the industry

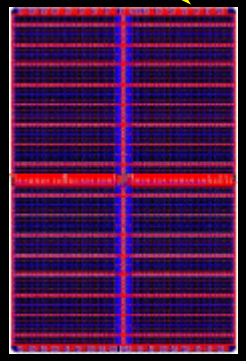
FLASH Extension	FRAM	MRAM	PCRAM	RRAM	PCM - SS Electrolyte	Polymer/ Organic	Mechanical	3D	Thyrister
Trap Storage	Ramtron	IBM	Ovonyx	IBM	Axon	Spansion	Nantero	Matrix	T-RAM
Saifun *NROM*	Fujitsu	Infineon	BAE	Sharp	Infineon	Samsung	STMicro	(Sandisk)	Sony
Tower	STMicro	Freescale	Intel	Unity		TFE	Hitachi	3D-ROM	
Spansion	TI	Philips	STMicro	Spansion		MEC		Samsung	
Infineon	Toshiba	STMicro	Samsung	Samsung		Zettacore		Macronix	
Macronix	Infineon	HP	Elpida			Roltronics		Infineon	
Samsung	Samsung	NVE	IBM			Nanolayer			
Toshiba	NEC	Honeywell	Macronix						
Spansion	Hitachi	Toshiba	Infineon						
Macronix	Rohm	NEC	Hitachi						
NEC	HP	Sony	Philips						
Nano-x'tal	Cypress	Fujitsu							
Freescale	Matsushita	Renesas							
Matsushita	Oki	Samsung							
	Hynix	Hynix							
	Celis	TSMC							
	Fujitsu								
	Seiko Epson								

IBM working towards a 16GB part by 2010

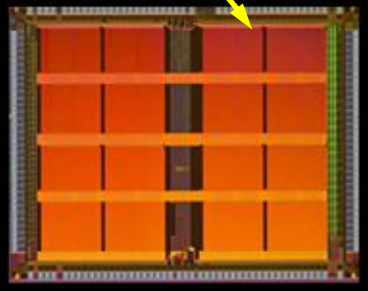
STMicroelectronics is claiming significant progress in the development of a new type of electronic memory that could eventually replace Flash memory technology



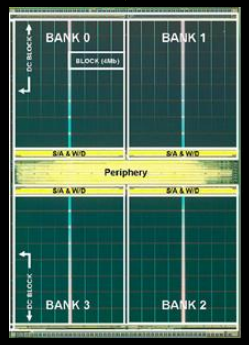
4Mb C-RAM (Product)
0.25um 3.3V



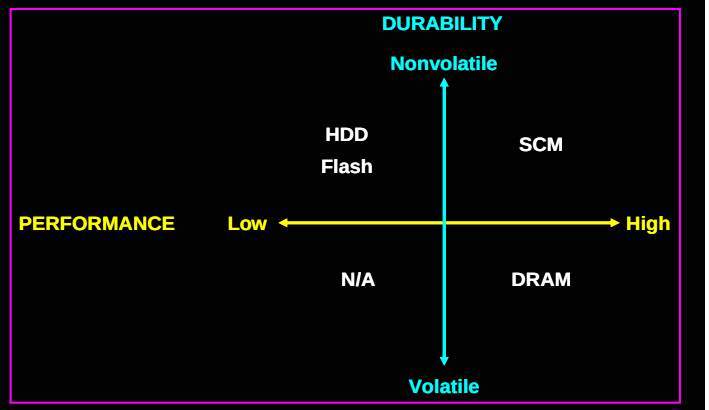
2Mb FRAM (Product)
0.35um 3.3V



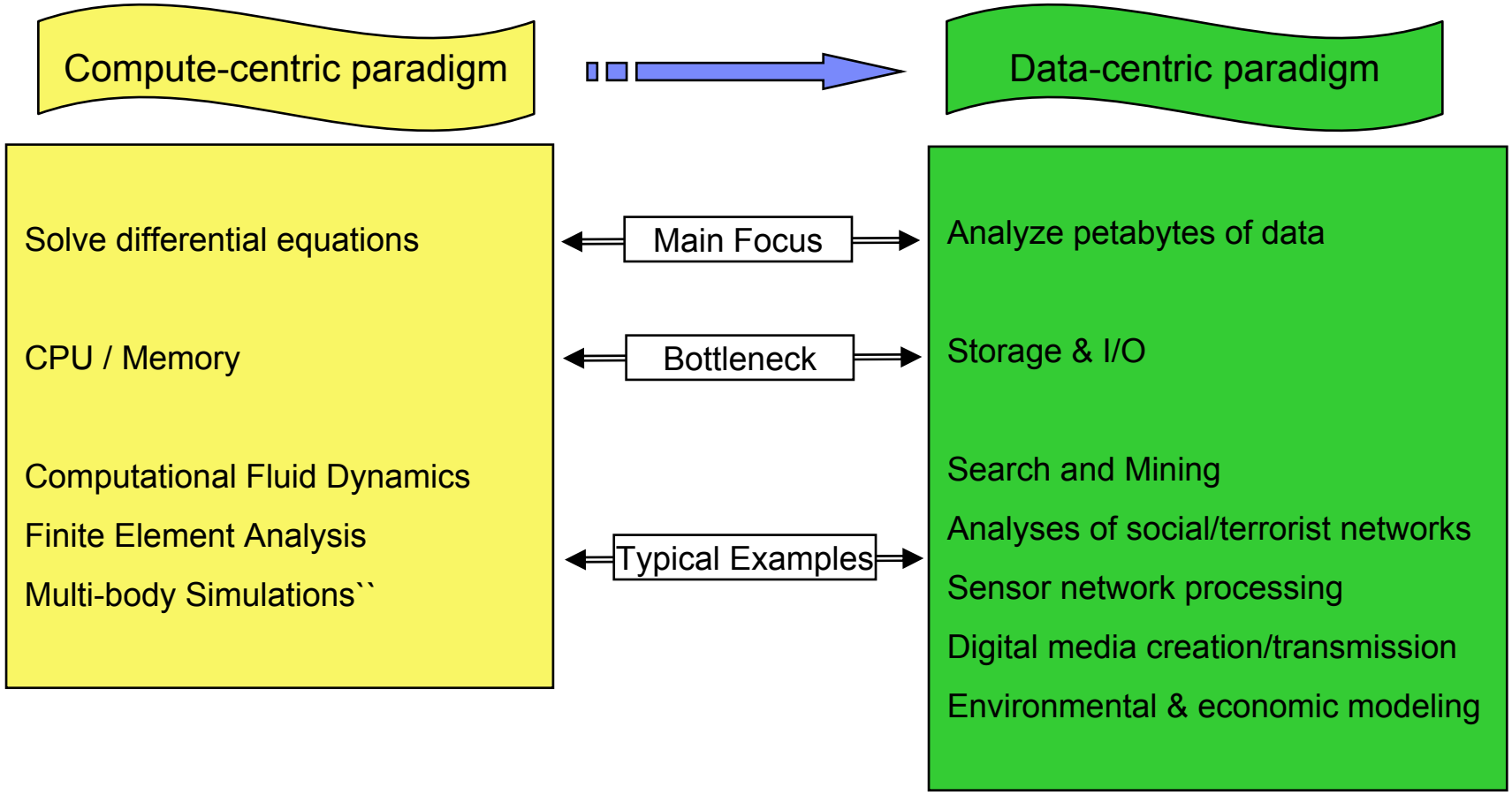
4Mb MRAM (Product)
0.18um 3.3V



512Mb PRAM (Prototype)
0.1um 1.8V



Critical applications are undergoing a paradigm shift

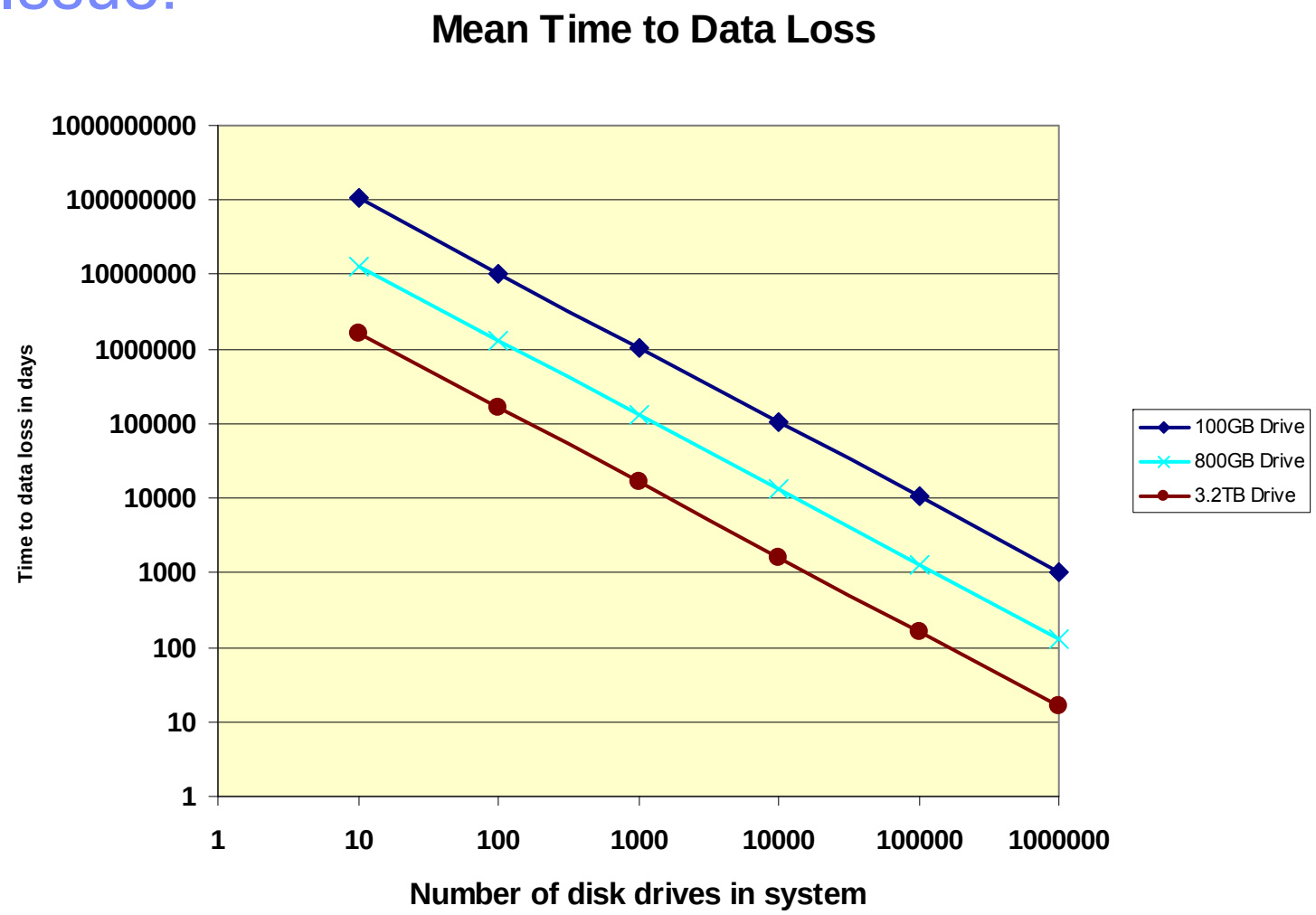


Thesis: *Disks or Flash can't keep up w/data centric applications*
Proposal: *Develop device technology and build a high density array and demonstrate performance and endurance for the data-centric paradigm*

What are the limitations with disks?

- **Bandwidth – Access Time – Reliability - Power**
- **Disk Performance improves very slowly**
  - Gap between processor and disk performance widens rapidly
  - Bandwidth gap can be solved with many parallel disks
    - but need 10,000 disks today, 500,000 disks by 2020
      - *but that's just for a traditional high-end HPC system*
      - *data intensive problems are much worse*
  - Access time gap has no good solution
    - disk access times decrease only 5% per year
    - complex caching or task switching schemes help - sometimes
- **Newest disk generations are *less* reliable than older ones**
  - Data losses occur in even the best enterprise-class storage systems
- **Disk power dissipation is a major factor in data-centric systems**

HDD Issue:



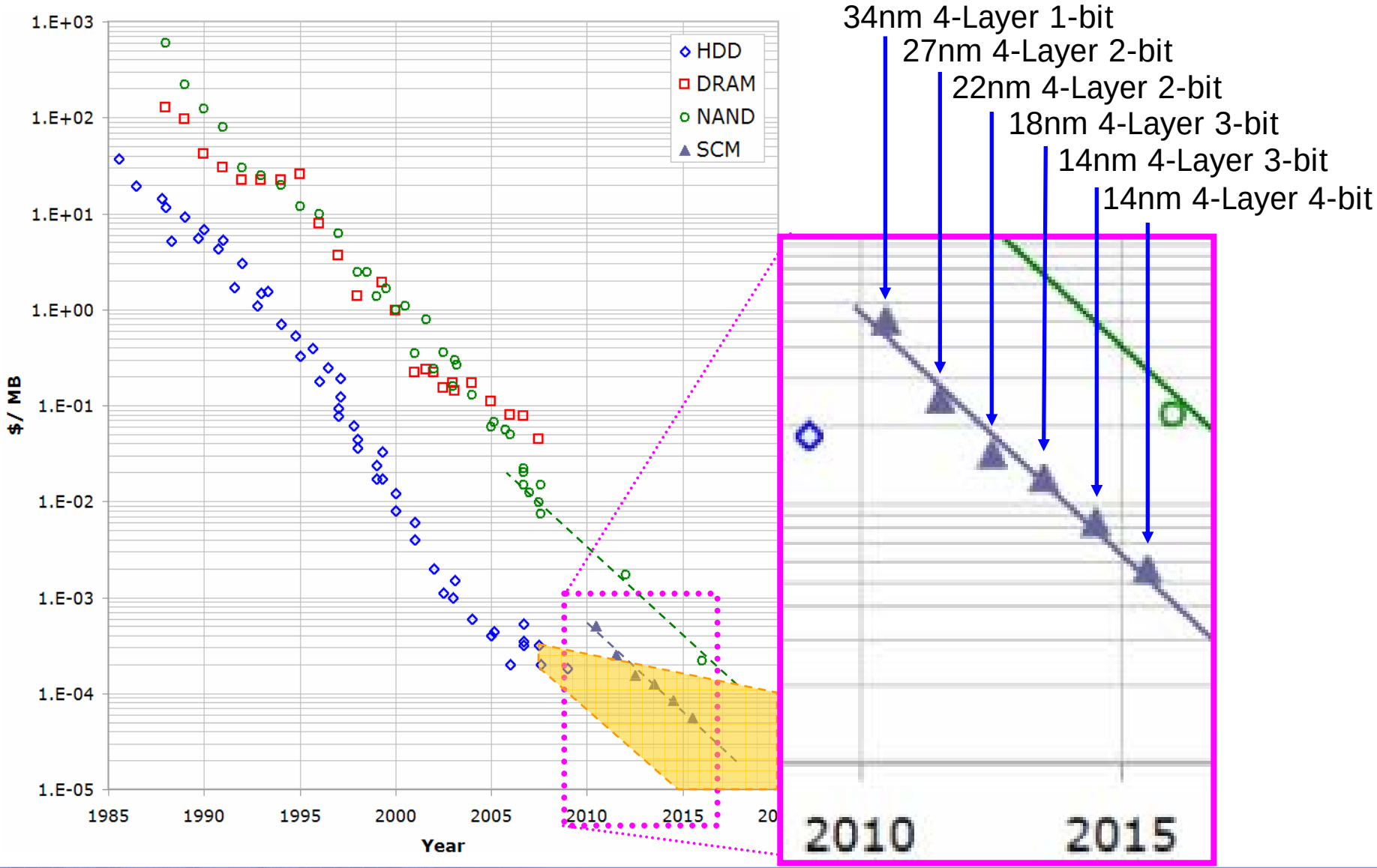
Assumptions: MTBF 1 Million Hours and RAID5

What are the limitations with Flash?

- Read/Write Access Times – Write endurance – Block architecture

- Flash Performance showing no improvement
  - Gap between processor and Flash performance continues to widen
  - Write endurance $<10^6$ and showing no improvement trends
    - Need $>10^9$ to cater to frequent writes as data continually flows into the system
      - *Tomorrow's hand-held devices will be continuously updated*
      - *Intel applications characterized by continuous data streams*
  - Access time gap has no good solution

Storage Historic Price Trend and Forecast



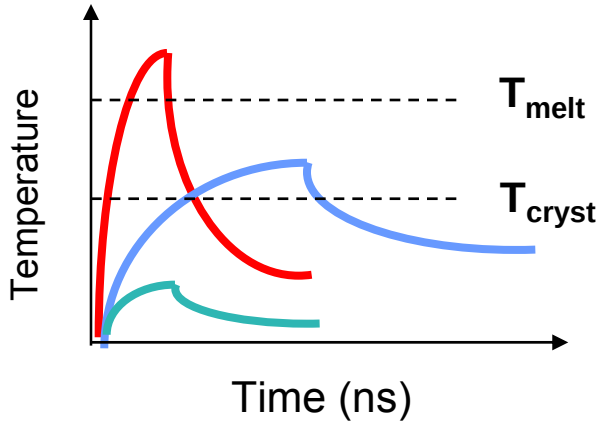
Storage Class Memory Target Specifications

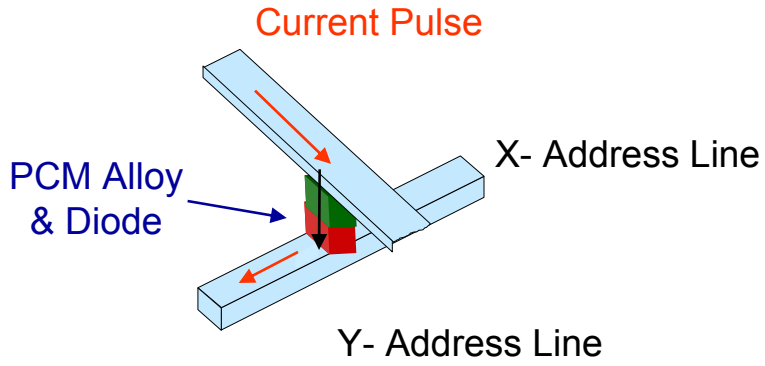
Access Time	**~100-200 ns**
Data Rate (MB/s)	**100**
Endurance	**10^9 - 10^{12}**
HER (/TB)	**10^{-4}**
MTBF (MH)	**2**
On Power (mW)	**100**
Standby (mW)	**1**
Cost ($/GB)	**<5.5**
CGR	**35%**

Very challenging to achieve in combination

SCM Basic Concepts (Phase Change Example)

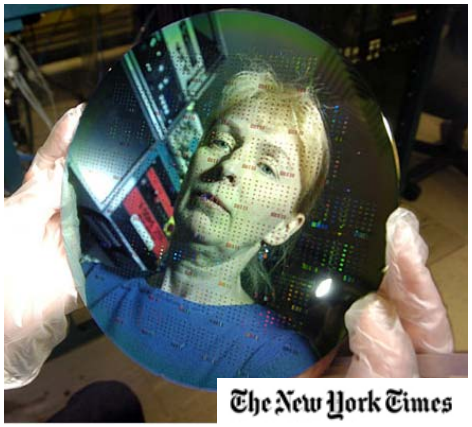
- **Using a phase transition of a *Ge-Sb-Te alloy to store a bit***
- ***Ge-Sb-Te* exists in a stable amorphous and a stable crystalline phase**
  - Phases have very different electrical resistances
- **Transition between phases by controlled heating/cooling**
  - Write '1' : short (10ns) intense current pulse melts alloy crystal => amorphous
  - Write '0' : longer (50ns) weaker current pulse re-crystalizes alloy => crystalline
  - Read : short weak pulse senses resistance, but doesn't change phase
- **Non-Si based proprietary diode materials being developed for high-ON current density (> 10^7 A/cm^2 – needed for PCM) and ultra-low OFF current density (< 1 A/cm^2).**



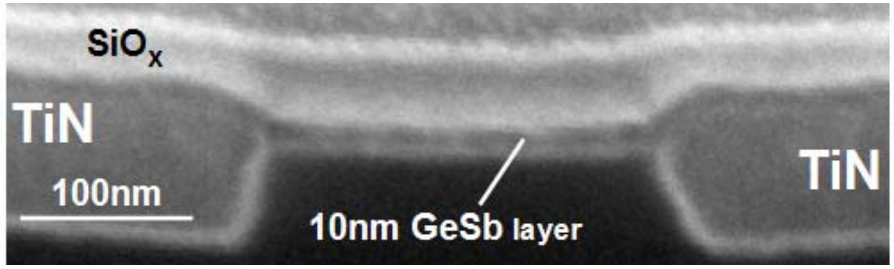


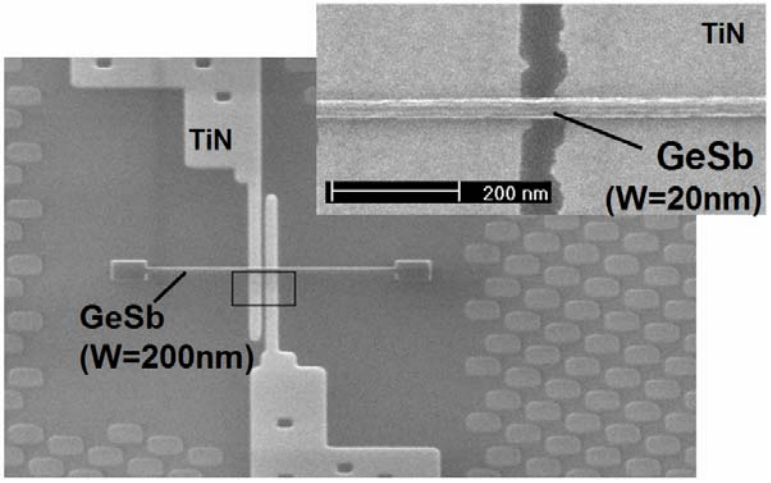
Phase-Change Nano-Bridge

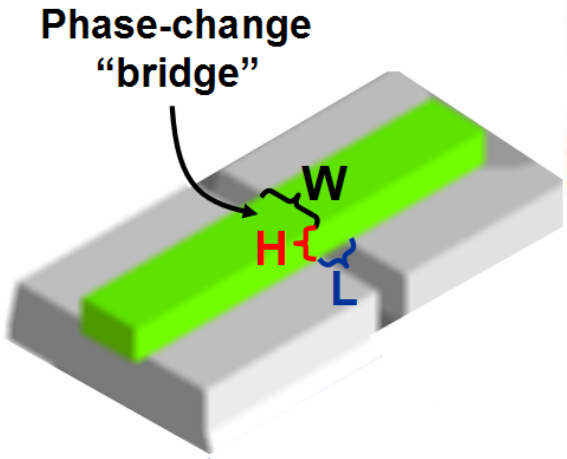
- Prototype memory device with ultra-thin (**3nm**) films demonstrated Dec '06



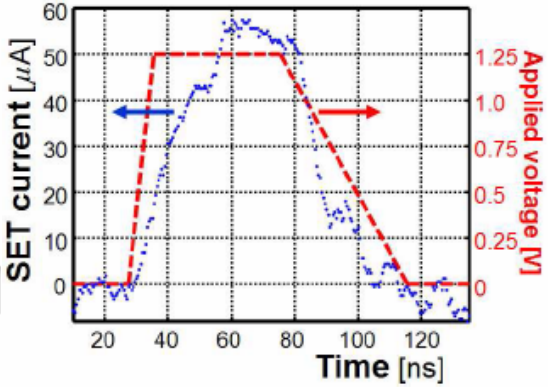
- 3nm * 20nm → 60nm^2
  ≈ Flash roadmap for **2013**
  → **phase-change *scales***
- Fast (<100ns SET)
- Low current (< 100μA RESET)

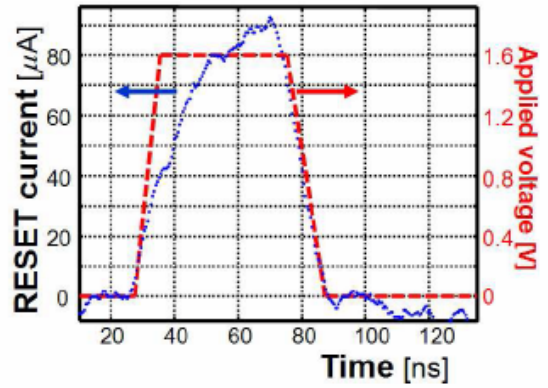




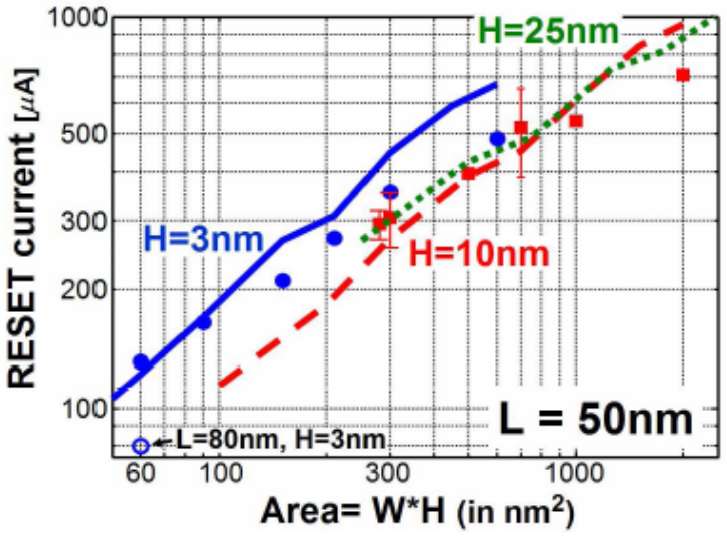


W defined by **lithography**
H by **thin-film deposition**



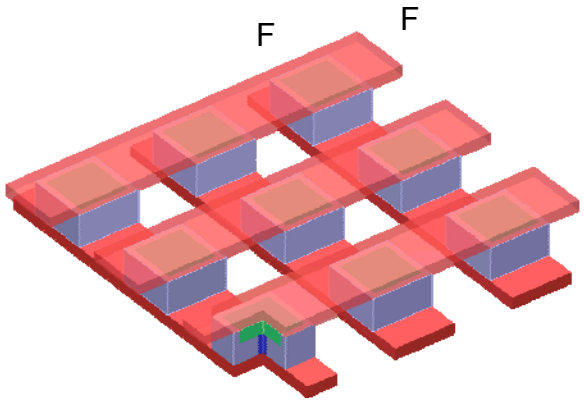


Current scales with area



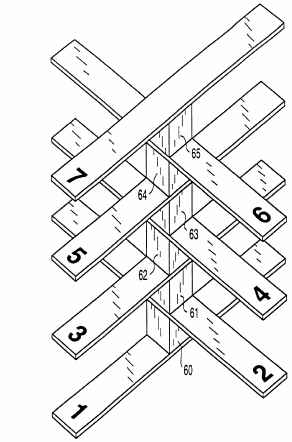
Processing Cost and F²

- **The bit cell size drives the cost of any memory**
- **Cell area is expressed in units of F² where F is the minimum lithographic feature of the densest process layer**
  - Half pitch dimension of metallization connecting drain and source for ICs
  - MR sensor width in magnetic recording
- **Cell areas**
  - DRAM 8F² ➔ 6F²
  - NAND 4F² ➔ 2F²
  - SRAM 100F²
  - MRAM 20F² -- 40F²
  - Hard Disk 0.5F² ➔ 1F²
  - ….



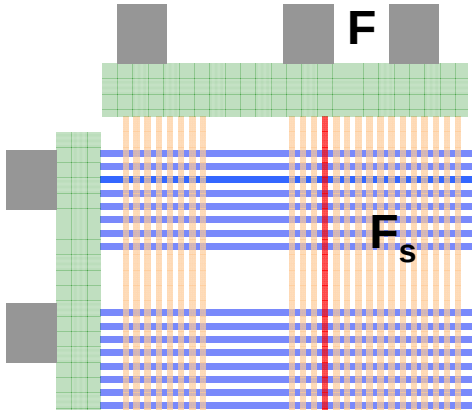
Low Cost Requires High Density

3D Lithographic Crosspoint Memory



$4F^2$ / (# of Layers)

2D Sub-Lithographic Crosspoint Memory



Blocks of Nanoscale Crossbar Arrays (NCA

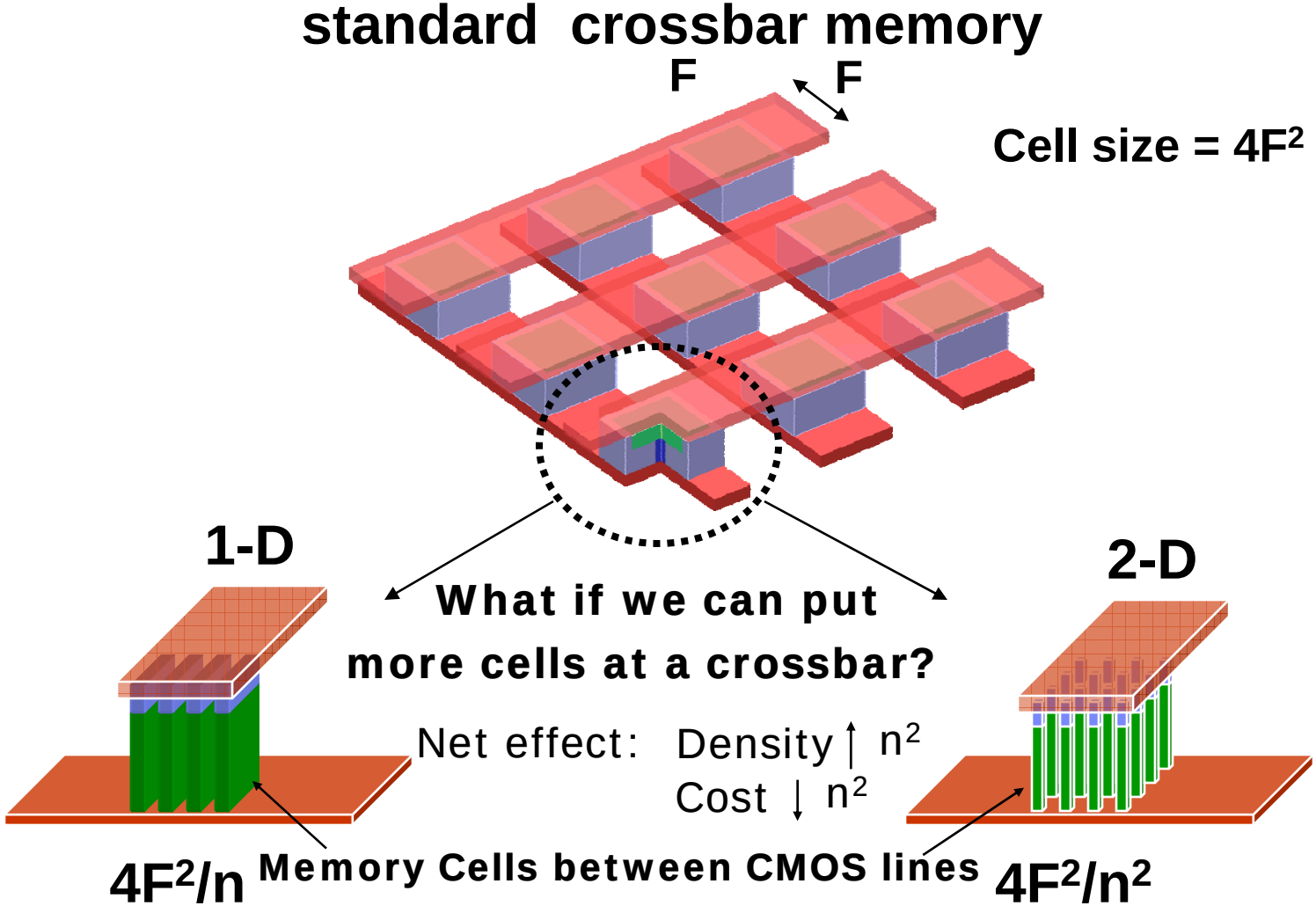
$4F^2$ / (# of Nodes)^2

Need Effective Cell Size < $4F^2$ (F → Lithography Half-Pitch)

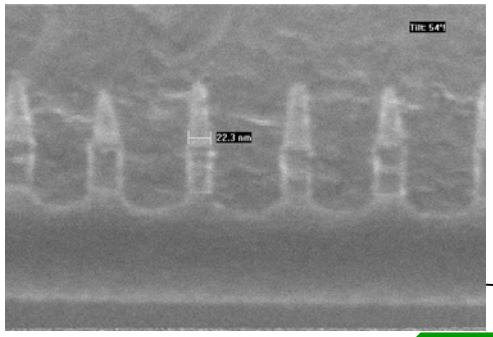
2D → Better Scaling & Fewer Process Steps, but Requires

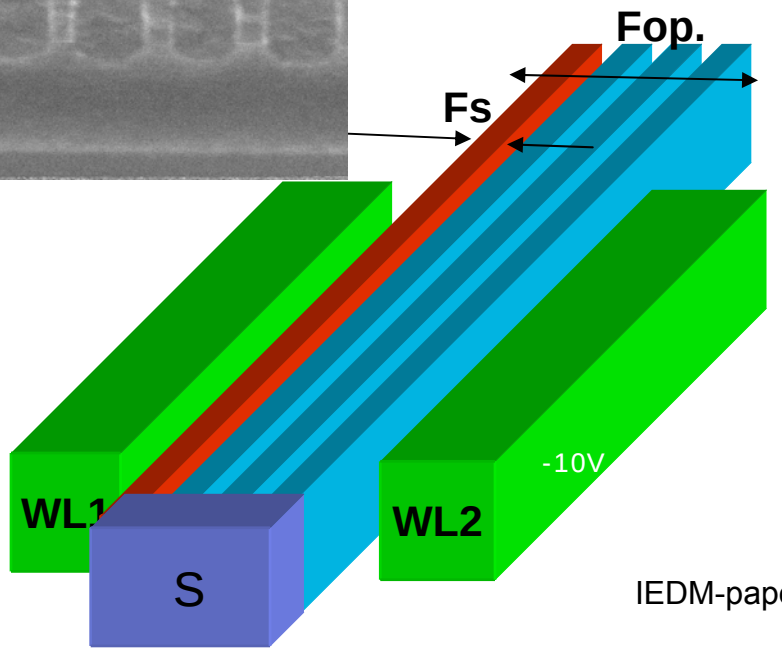
- Interface Between Litho (F) & Sub-Litho (Fs)
- Viable Method to Manufacture Sub-Lithographic Arrays

Crossbar Memory Fundamentals

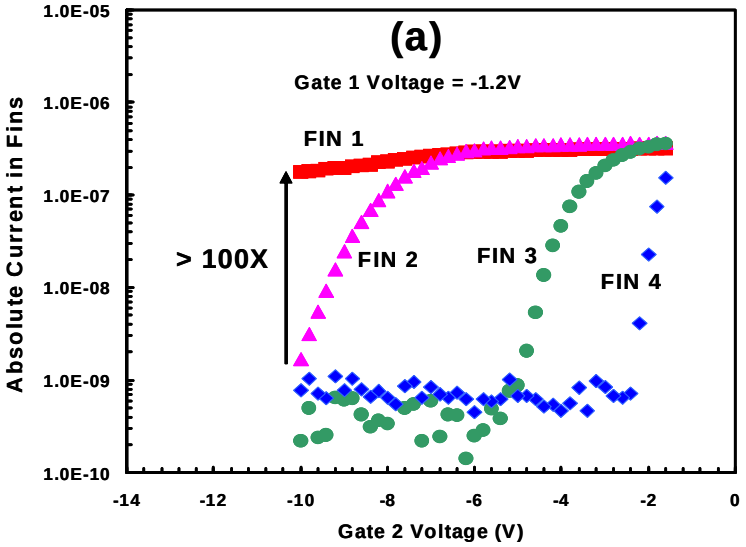


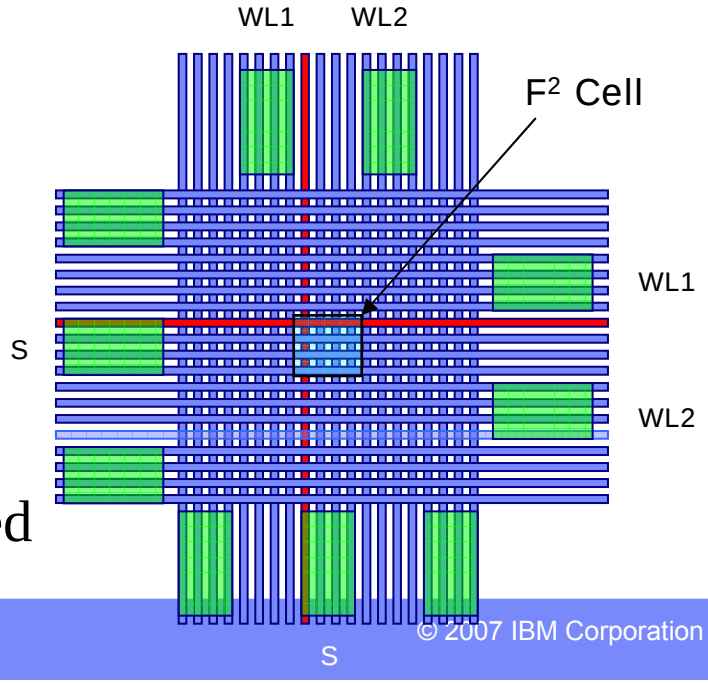
Micro-Nanoscale Decoder





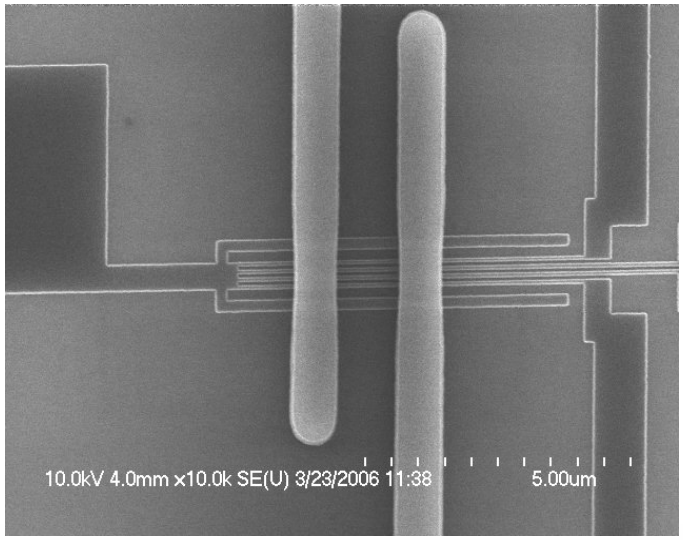
IEDM-paper 2005





- Sub lithographic feature is selected by moving depletion across the fine structure
- Modulating signal is brought in by lithographically defined lines
- Fins down to sub 20 nm have been addressed

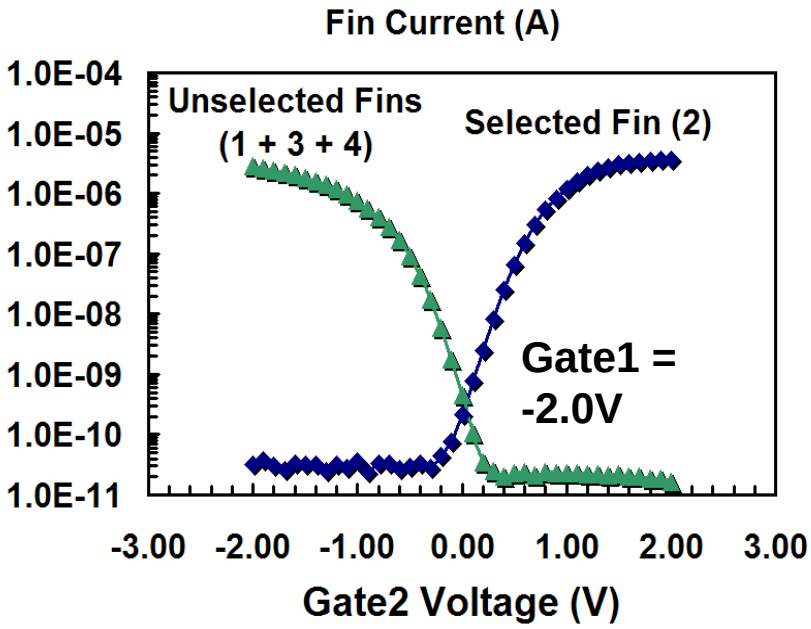
MNAB Concept Demonstrated

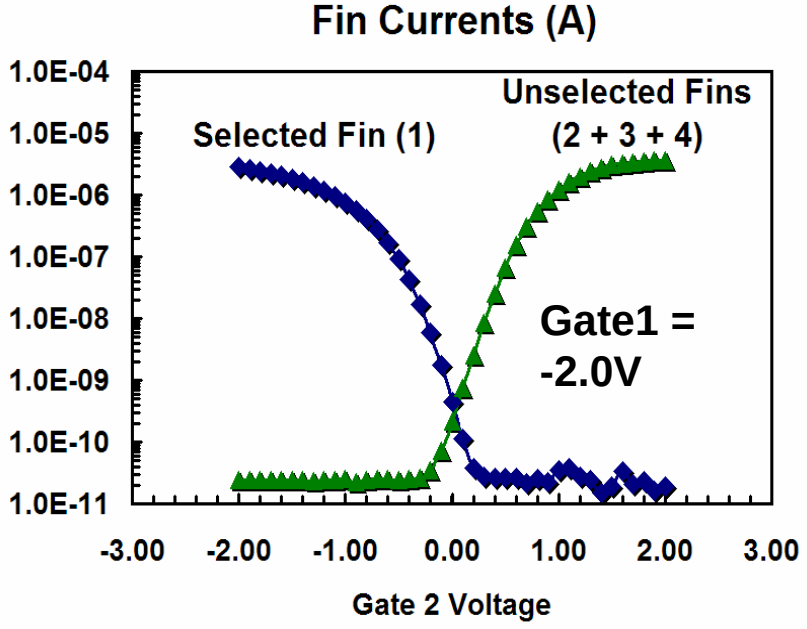


100nm Pitch MNAB Devices Fabricated by E-Beam Lithography

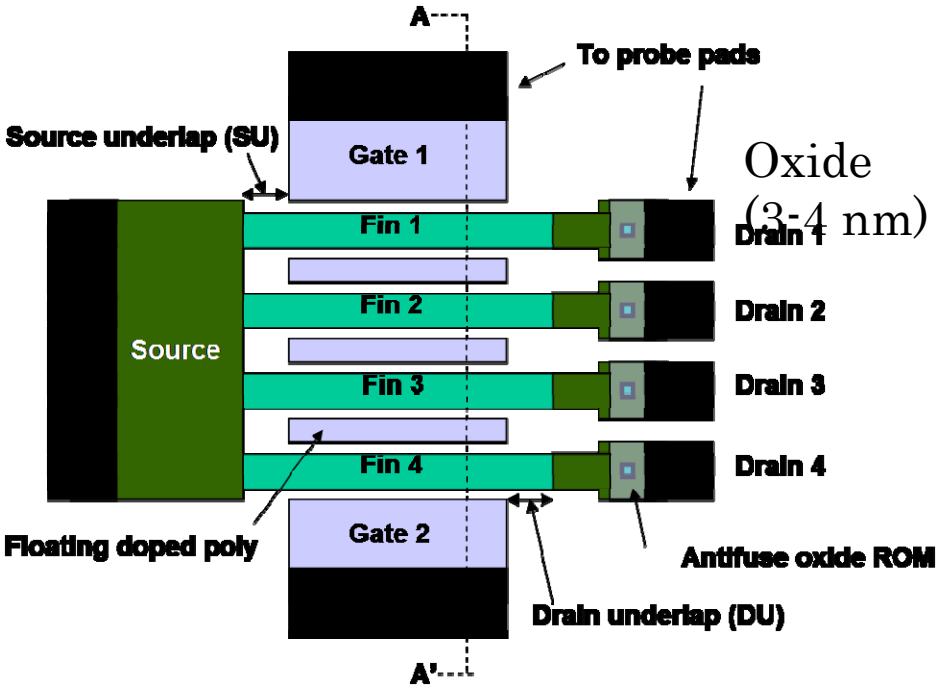
Obtained Fully Functional Devices

Selectivity > 10^5

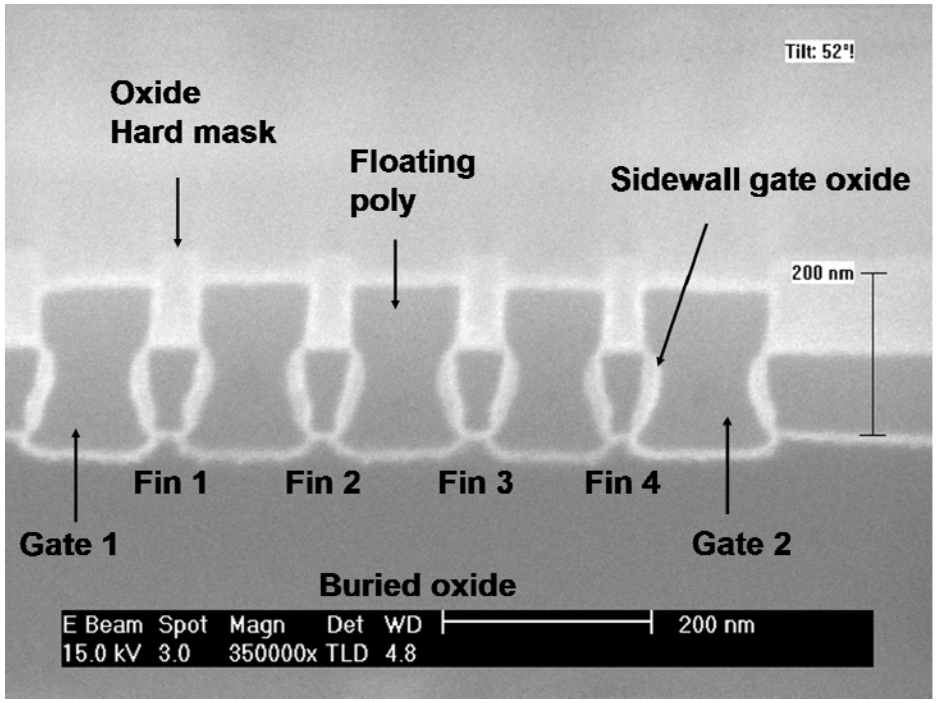




Combining Micro-Nano Decoder and ROM



4-fin UMB+ROM test structure



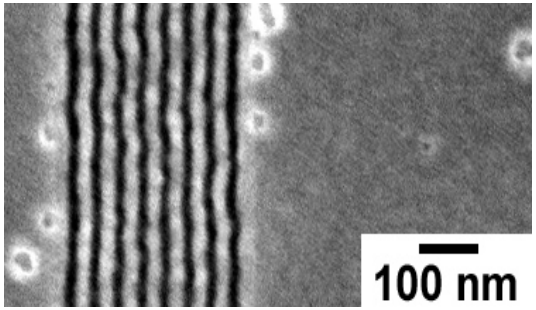
FIB x-SEM through gated fins (A-A')

✓ Successful integration of UMB with memory element
(2 terminal oxide antifuse ROM)

✓ Verified operation over all bit sequences for 4-fin UMB+ROM

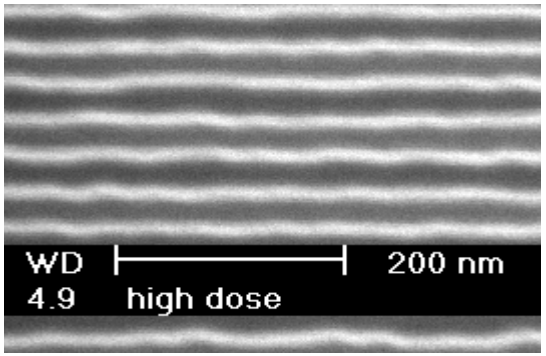
Nanoscale Patterning Techniques

Self Assembly



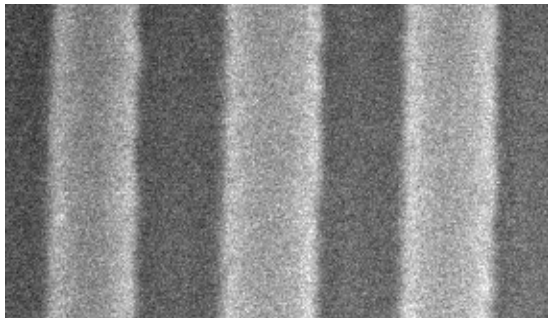
(IBM T J Watson Research Center)

Spacers



Frequency doubling – 40 nm to 20 nm pitch (IBM)

Nanoimprint Lithography



(IBM Almaden)

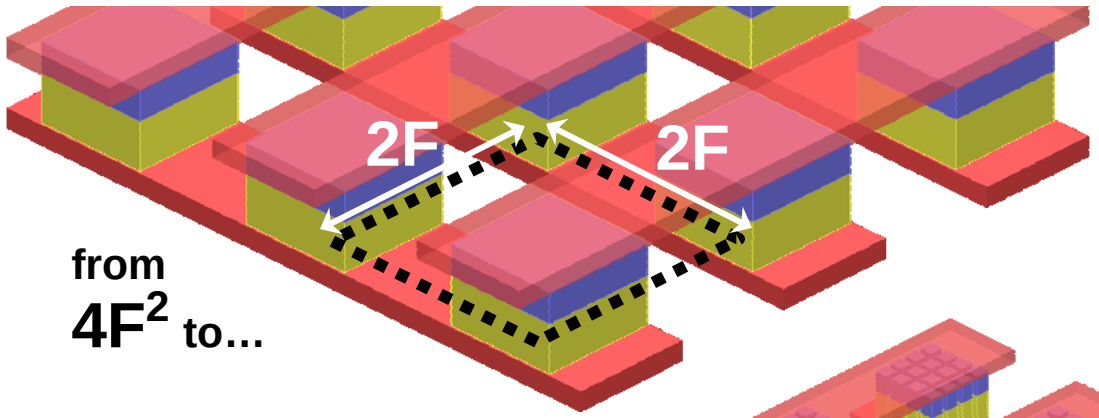
- **Various nanoscale patterning techniques exist.**
- **Sub 20 nm pitch demonstrated.**
- **Only regular line / space patterns possible.**

Paths to ultra-high density memory

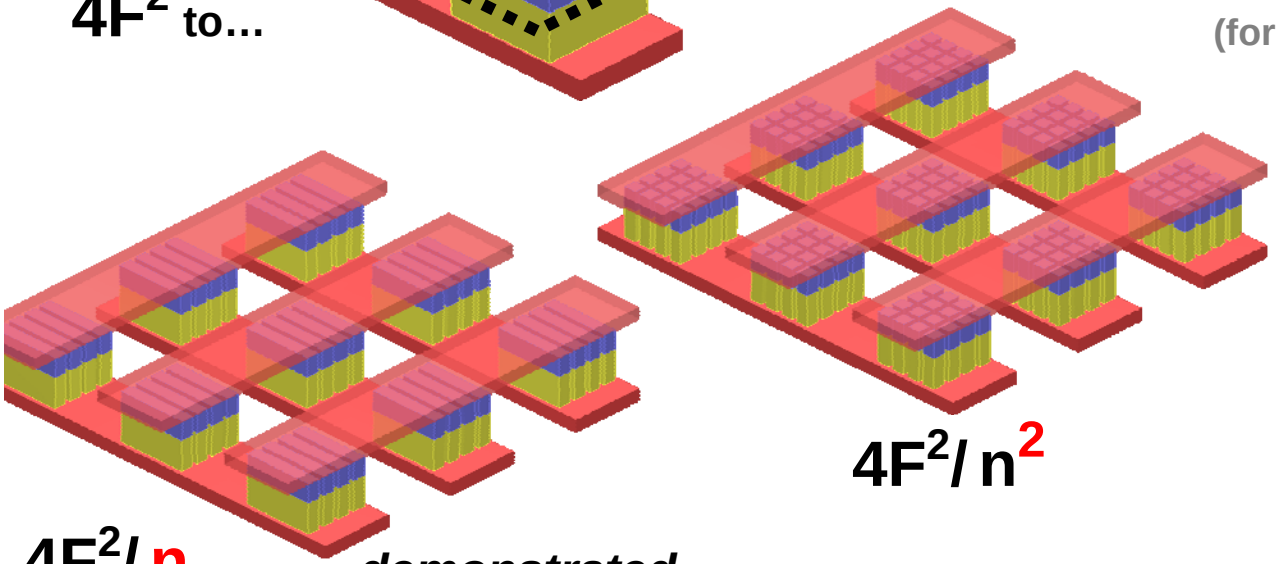
At the 32nm node in 2013, MLC NAND Flash is projected* to be at

if we could shrink $4F^2$ by	*density*	*product*
	43 Gb/cm² →	32GB
n	97 Gb/cm² →	64GB
n^2	387 Gb/cm² →	256GB
Ln^2	1550 Gb/cm² →	~1 TB

(for m=n=4)

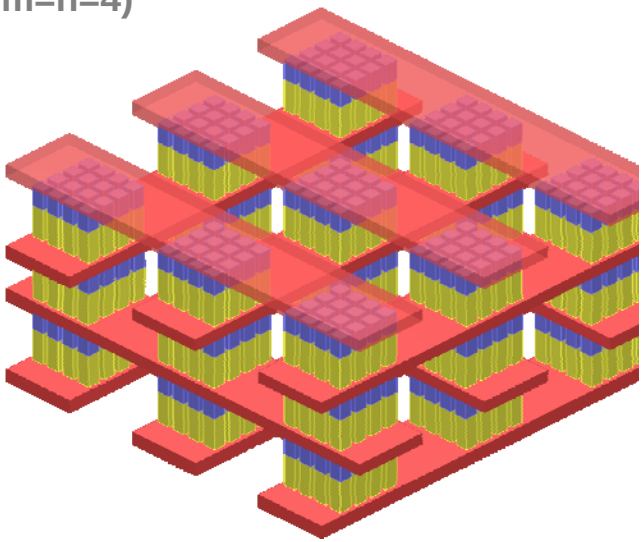


from $4F^2$ to…



$4F^2/n$ *demonstrated* (at IEDM 2005)

$4F^2/n^2$

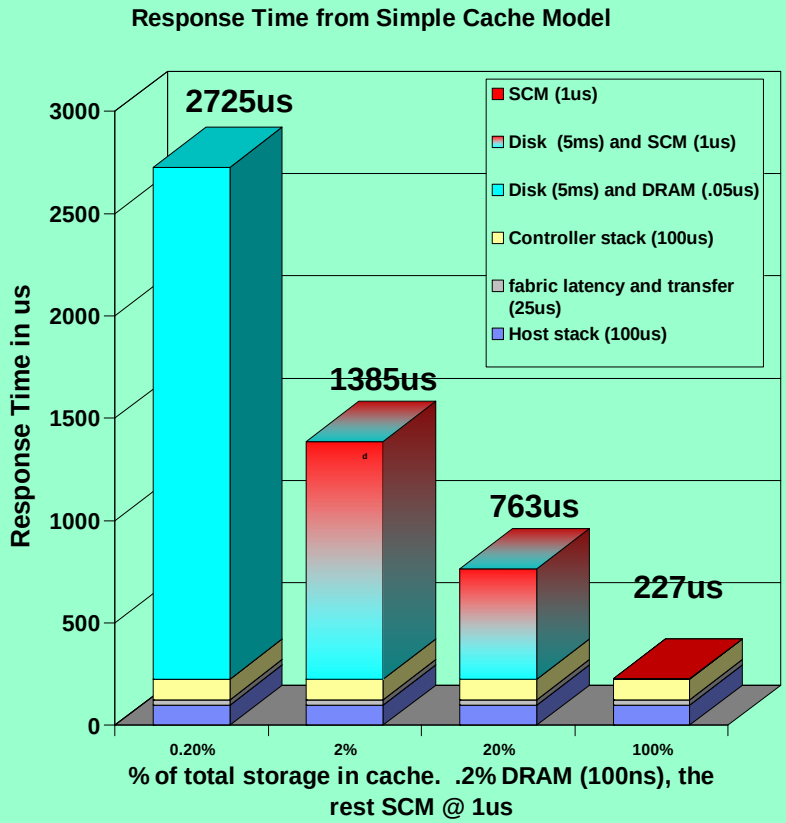


$4F^2/Ln^2$

* 2006 ITRS Roadmap

Value Proposition for SCM in Storage Controllers

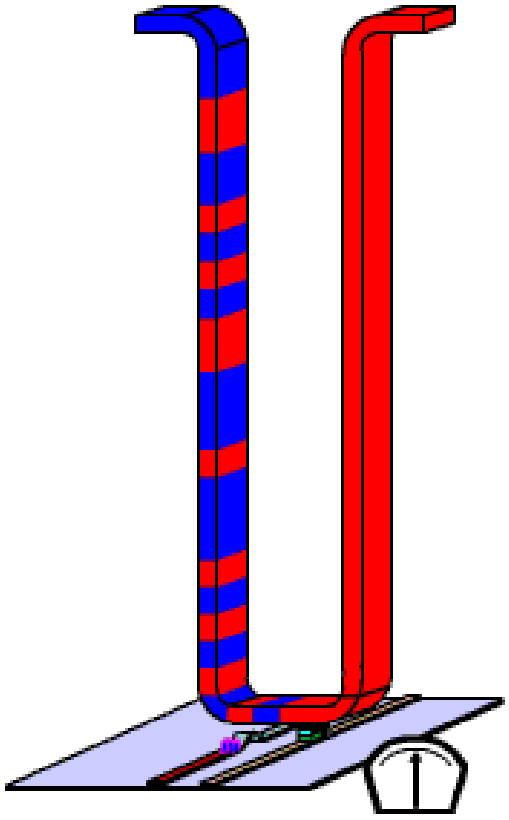
- **Significantly improved cost/performance long before SCM competes with DASD on cost**
- **Simple cache model (cube root rule), queuing effects ignored, miss rate starts at 50%**
- **Columns in figure**
  - 1st – business as usual (BAU), DRAM cache @ .2% of DASD capacity
  - 2nd – same cost as BAU, but DRAM replaced by SCM, SCM @ 2% of DASD capacity, perf. ~2x
  - 3rd – **Hierarchical storage** cost now 2.8x 1st column, SCM capacity now 20% of DASD, perf. >3.6x
  - 4th – cost now 1-10x, SCM only storage used in system. Perf. >12x
- **Performance assessed at application interface to OS. So, I/O stack in host, fabric latency and storage controller microcode processing time and data transfer time are included**



SCM impact on large HPC storage systems

- **Large file throughput**
  - >>10x improvement per TB of file system possible
    - Limited by interconnect and controller bandwidth
    - Limited by file system OS software overheads
  - Good match for check-pointing
  - Bulk storage costs high
- **Small file and metadata access rates**
  - Access rate improvement >100 feasible
  - Limited by software stack and controller overhead

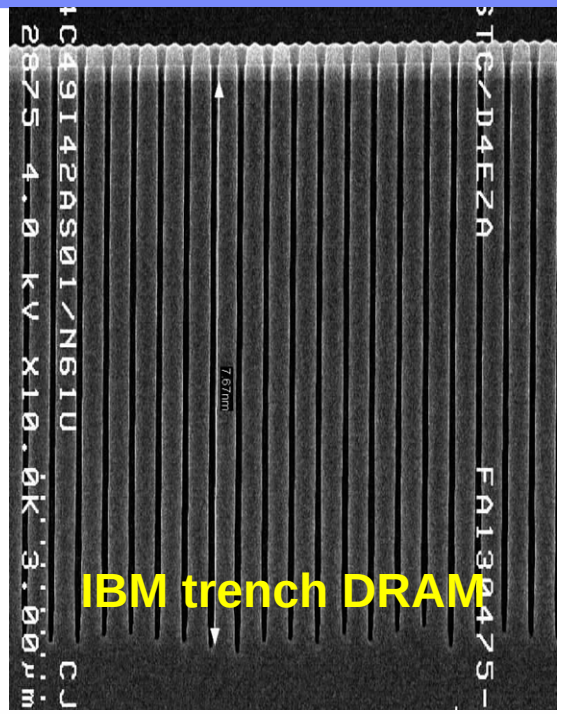
Magnetic Racetrack Memory

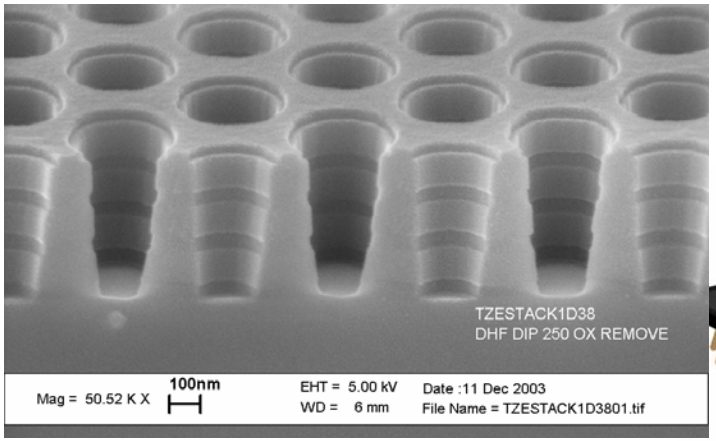


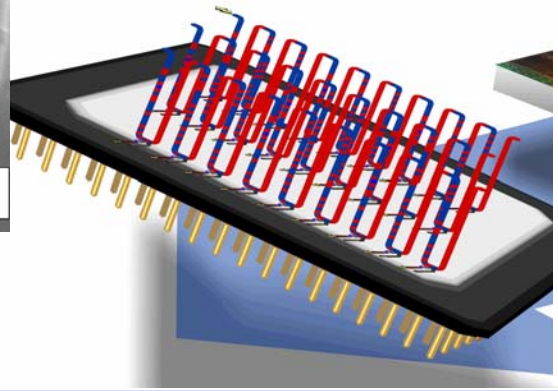
- Data stored as pattern of domains in long nanowire or "racetrack" of magnetic material.
- Data stored magnetically and is non-volatile.
- Current pulses move domains along racetrack – *no moving parts, just the patterns move.*
- Each memory location stores *an entire bit pattern* (10, 100, 1000 bits?) rather than just a single bit.

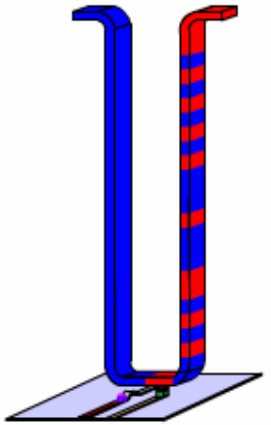
Magnetic Race-Track Memory

- **Information stored as domain walls in vertical "race track"**
  - Data stored in the third dimension in tall columns of magnetic material
- **Domains moved around track using nanosecond pulses of current**
- **10 to 100 times the storage capacity of conventional solid state memory**



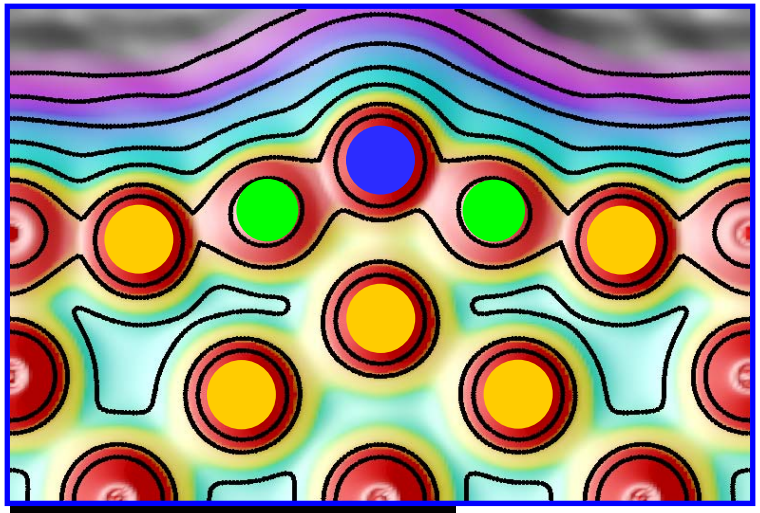


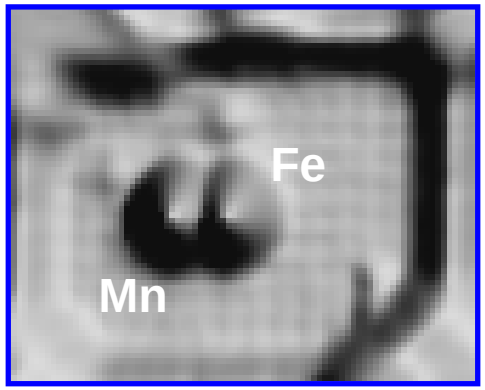


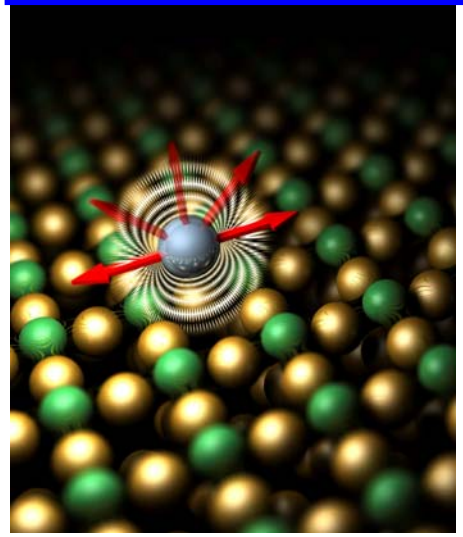


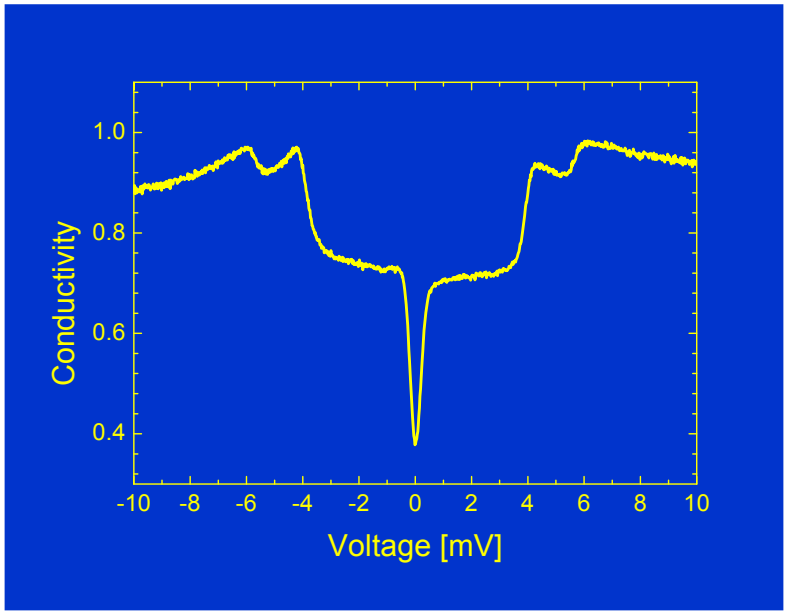
Magnetic Race Track Memory
S. Parkin (IBM), *US patents*
6,834,005 (2004) & 6,898,132 (2005)

Large Magnetic Anisotropy for Single Atoms









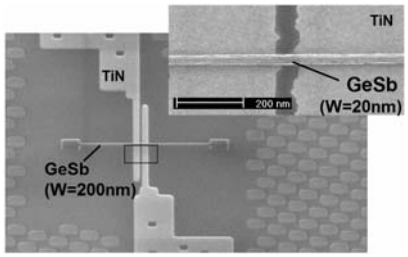
- The energy that is required to change the direction of a single spin was measured.
- Large single-atom magnetic anisotropy for **iron of about 6 meV.**
- About 50x weaker anisotropy for manganese on same surface.
- Spin excitation spectroscopy reveals spin energy levels, including their magnetic field dependence.
- DFT calculations elucidate surface structure and leads to same total spin as experiment.
- **GOAL**: engineer very large magnetic anisotropy to demonstrate data storage.

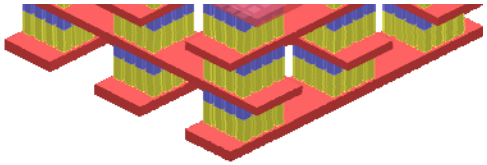
The Future of Memory?

Storage Class Memory:
solid-state non-volatile memory
at hard-drive prices

- **Phase-change memory** – low cost because >1 bit / $4F^2$

~2013?

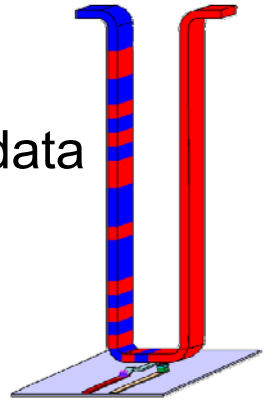




- **Racetrack memory**

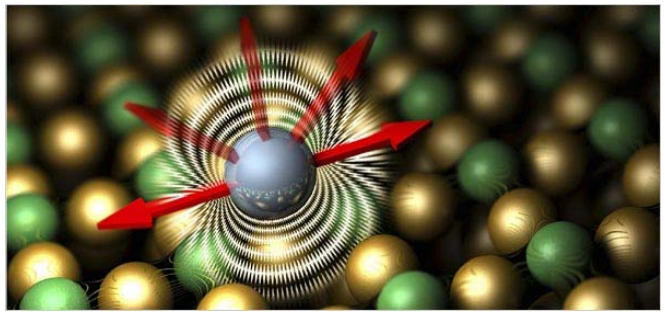
~2018?

– a 3-D nano-warehouse for data



- **Atomic memory** – "there's a lot of room at the bottom…"

~2030?



Science & Technology
IBM Almaden

Innovation and Impact

- **Storage class memory (SCM)**
  - New nonvolatile solid state memory with fast access and high throughput
  - Robustness, volumetric density and power significantly better than disks
- **Will revolutionize memory/storage hierarchy**
  - >10x throughput, >100x transaction rate potential
  - Applications will be revamped to exploit new technology
- **New applications or significantly extended applications, e.g.**
  - Sensor/actuator systems with storage at the network edge
  - Mobile applications, e.g. semiautonomous video gatherers
  - …