Emerging Memories: Are They Energy Efficient Enough?

H.-S. Philip Wong
Stanford University

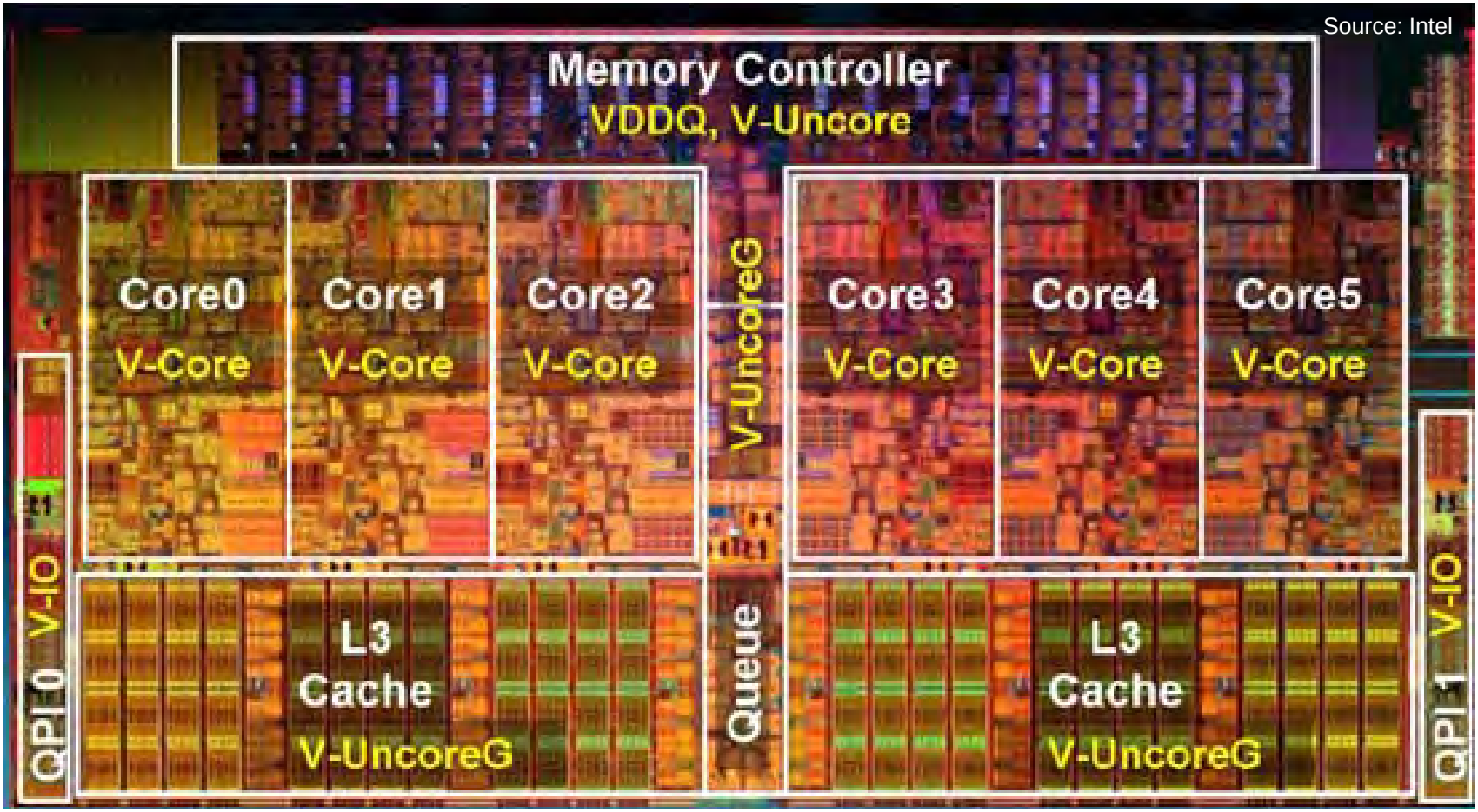
Memory – Key Enabler for New Applications

256GB

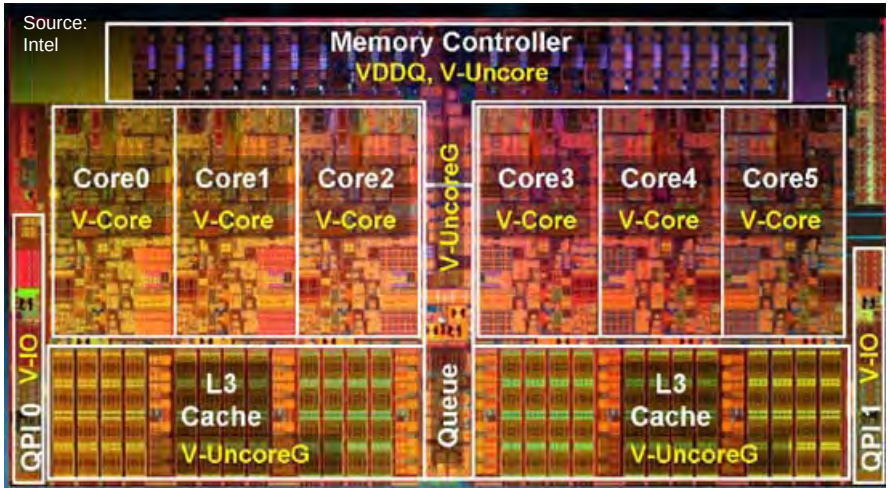
8GB

64GB

16GB


Source: Intel
Memory Controller
VDDQ, V-Uncore
Core0
V-Core
Core1
V-Core
Core2
V-Core
V-UncoreG
Core3
V-Core
Core4
V-Core
Core5
V-Core
QPI 0
V-IO
L3
Cache
V-UncoreG
Queue
L3
Cache
V-UncoreG
QPI 1
V-IO

Energy, Environment, Sustainability

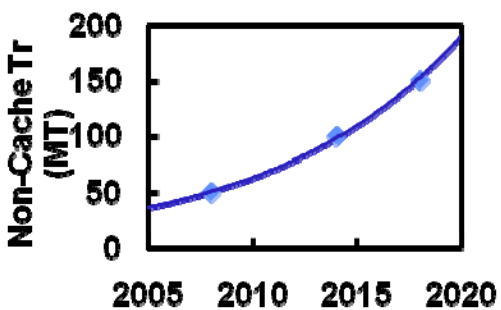


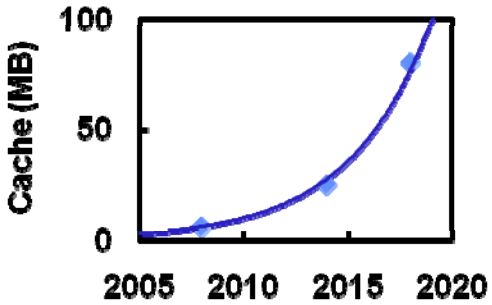
- **CPU consumes**
  - 27 – 57% of server power (2005 – 2007)
  - ~50% of CPU power consumed in devices
  - Exceeds total power from solar (US, 2007)

D.A.B. Miller, *Proc. IEEE,* p. 1166 (2009).

Compute and Memory Must be Balanced

- **Amdahl's system balance rules: "*the MB/MIPS ratio (α) in a balanced system is 1.*"**
  - Today's system: $\alpha \approx 4$ (and rising) [1]





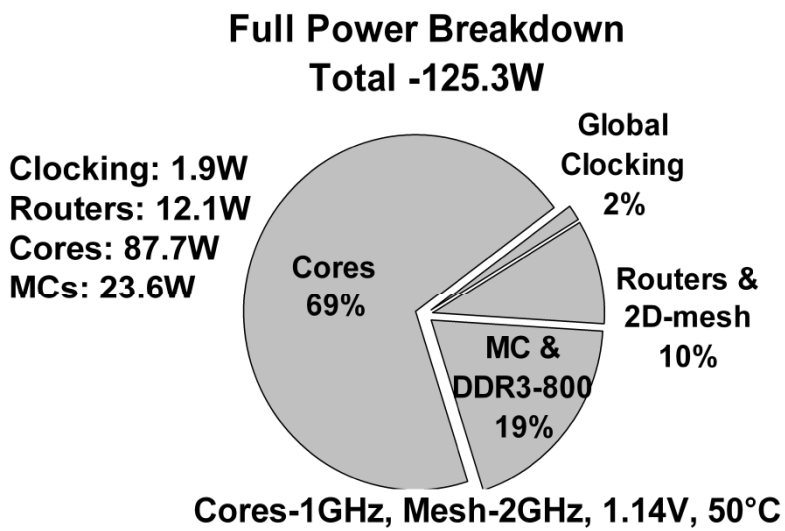
- **Non-memory transistors increase only 3X in 10 years [2]**
  - That's all you can afford (power)
- **Memory integration capacity will outpace logic > 10X**

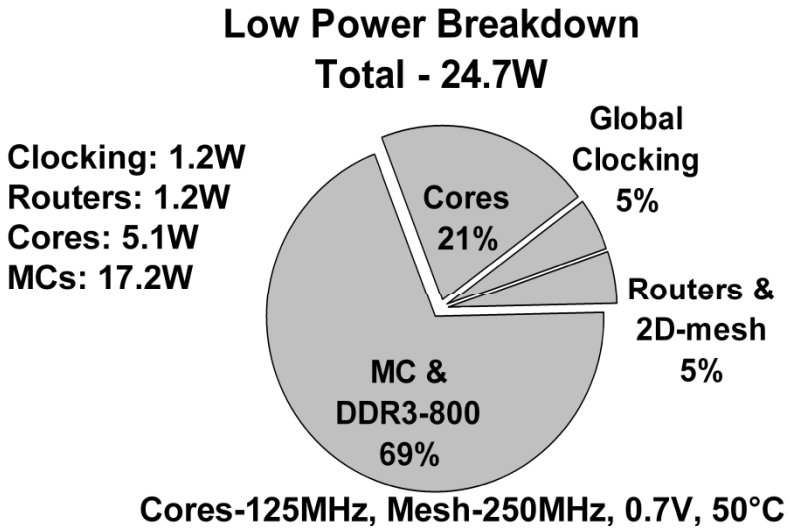
[1] J. Gray and P. Shenoy, "Rules of thumb in data engineering," in *Proc. Int. Conf. Data Engineering* (ICDE'00), 2000.

[2] S. Borkar, "The Exascale Challenge," *VLSI-TSA* (2010).

Demands from Above

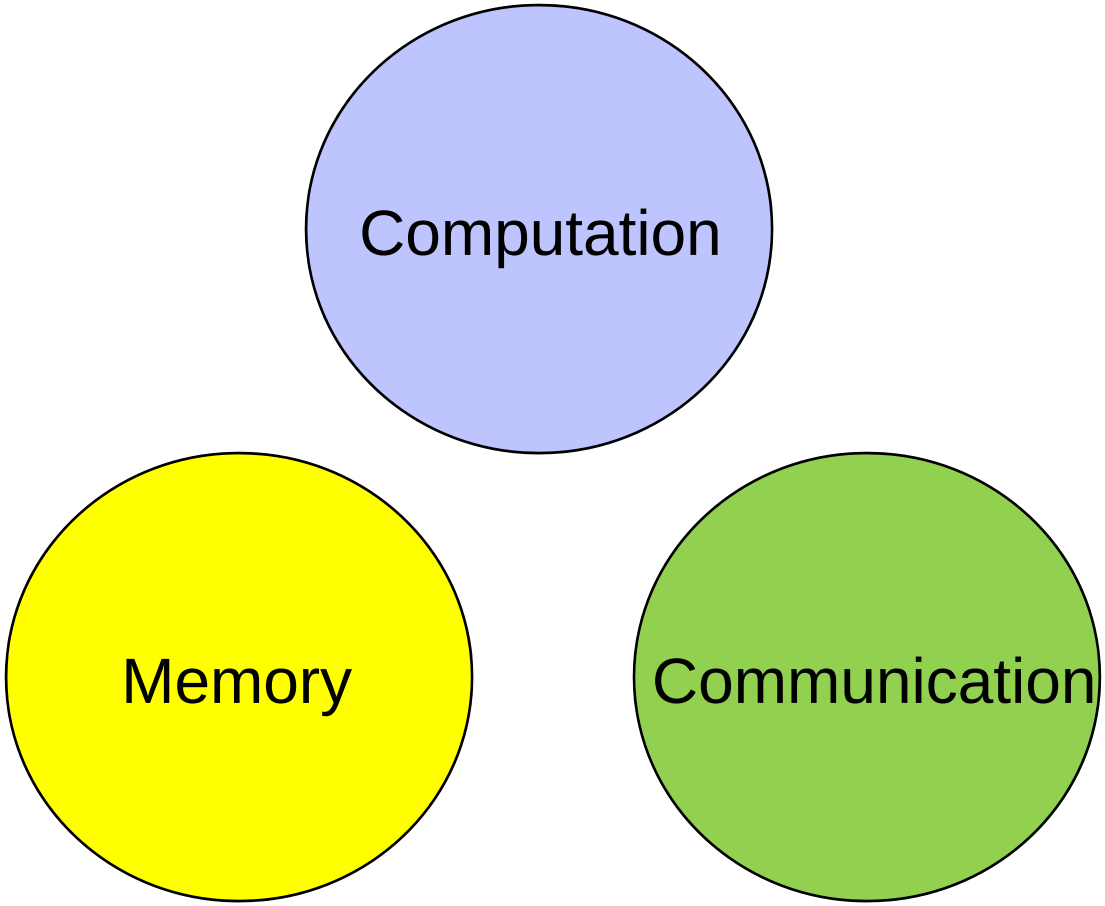
- **Power breakdown of Intel IA-32, 48-Core Processor in 45 nm CMOS – memory is a big piece**



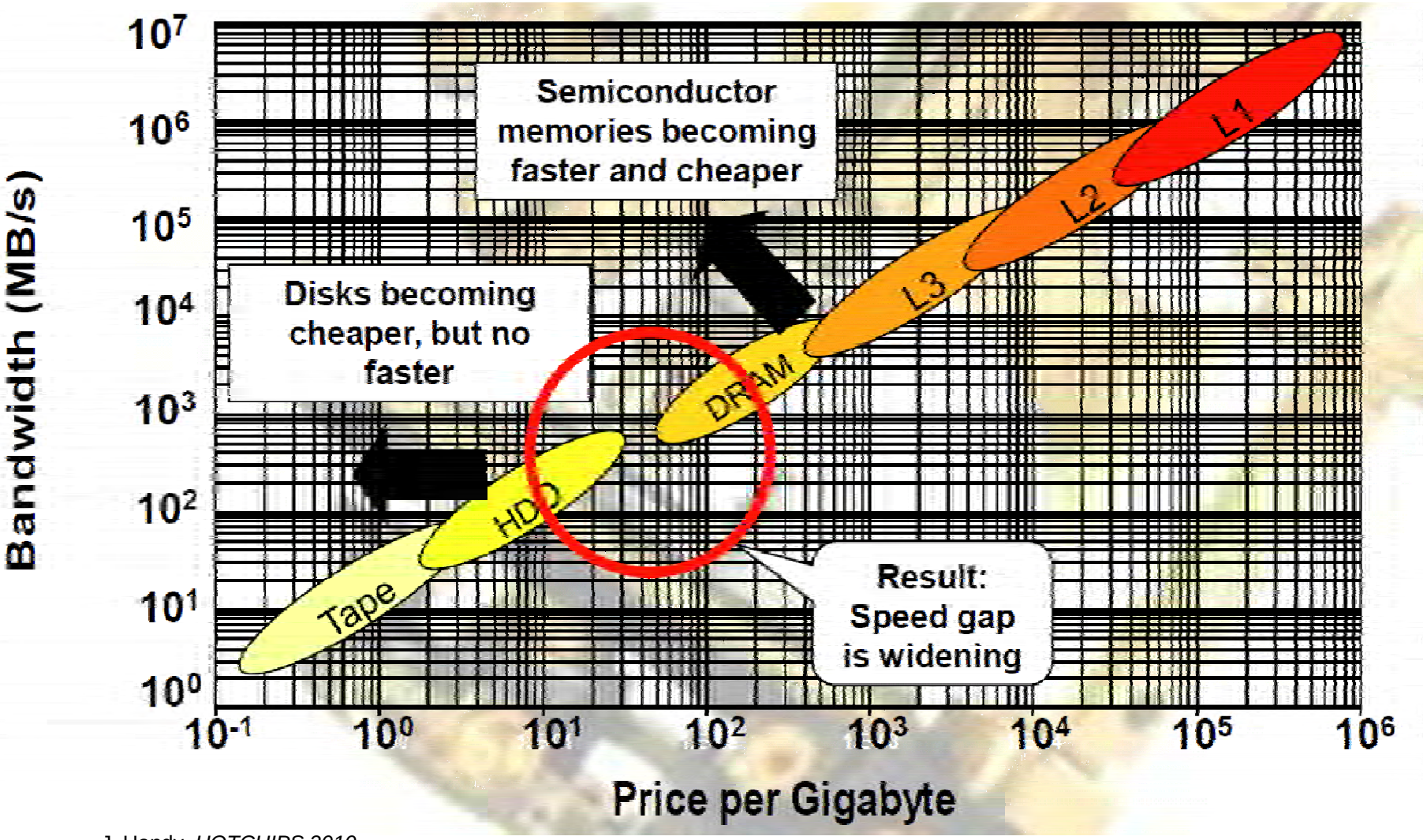


J. Howard, "A 48-Core IA-32 Message-Passing Processor with DVFS in 45nm CMOS," *ISSCC*, p. 108 (2010).

Nanoelectronics – Don’t Forget the Memory



DRAM/HDD Bandwidth Gap



J. Handy, *HOTCHIPS 2010*.

Opportunities For Memories

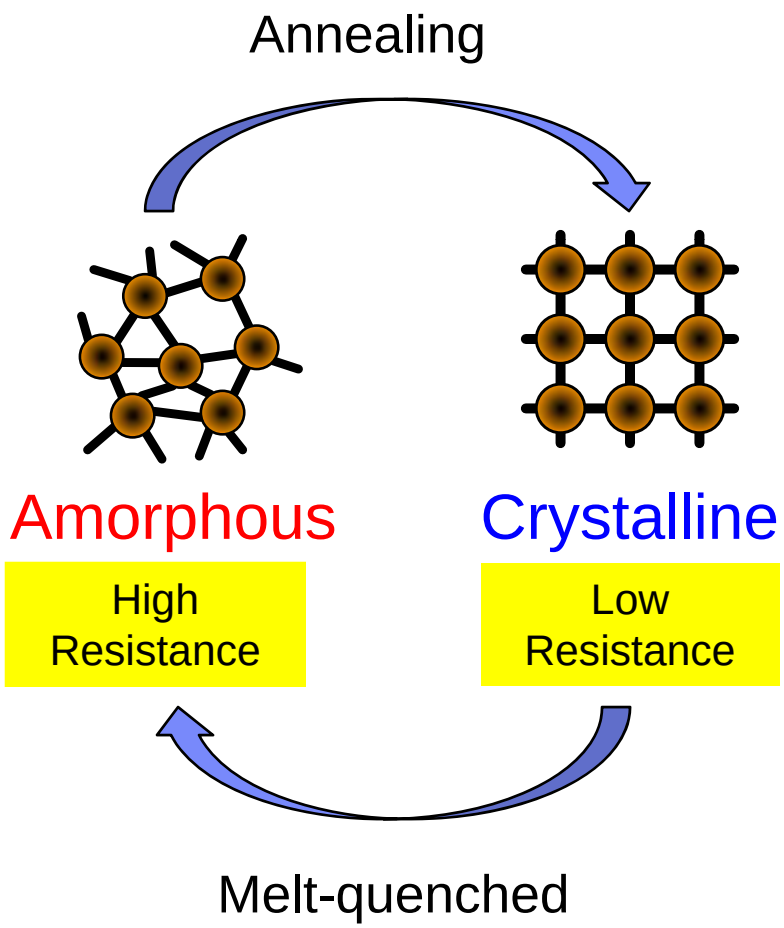
- **White space has opened up in memory hierarchy**
  - Speed/capacity gap
- **Lots of GB needed for various applications**
  - Need high density
- **Question is:**

Will the memories be energy efficient enough?

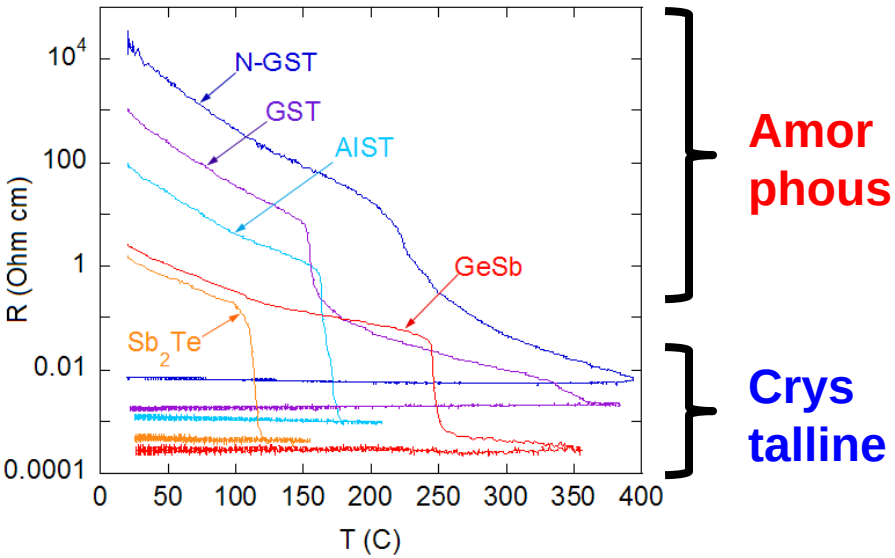
Emerging Memory Candidates

- **Phase change memory (PCM, PRAM, PCRAM)**
- **Spin torque transfer RAM (STTRAM)**
- **Resistive switching metal oxide RAM (RRAM)**
- **Conductive bridge RAM (CBRAM)**
- **…and many others that fall into the category**
  - I cannot make this into a good logic switch, but it has some hysteresis, so let's call that a memory

Principles of Phase Change Memory (PCM)



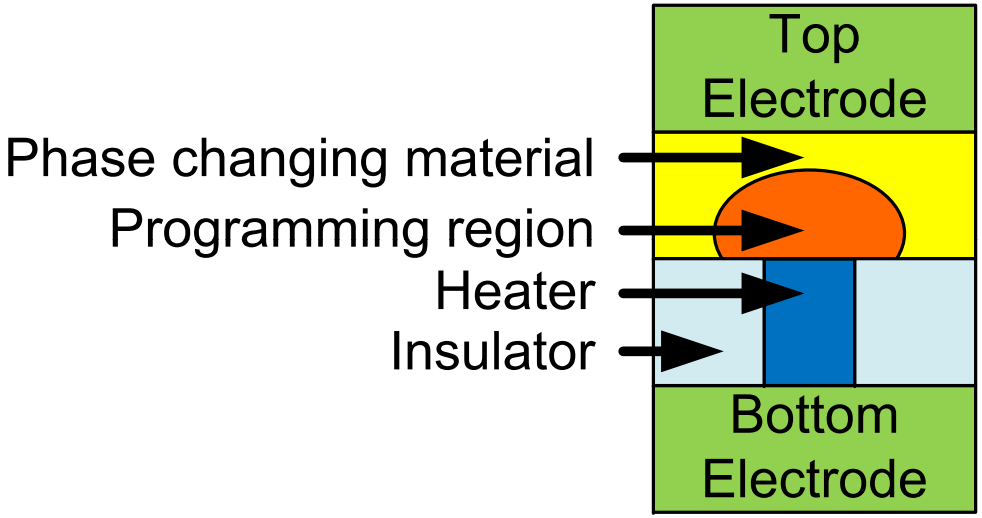
[Various phase-change materials]
$Ge_2Sb_2Te_5$ (GST), AIST (AgInSbTe), GeSb, Sb_2Te and etc..

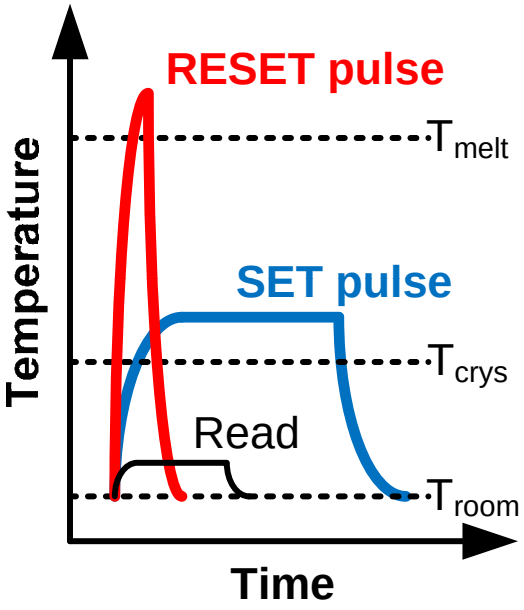


S. Raoux et al., *JAP*, v. 102, p. 094305, 2007.

- **<u>Resistance</u> change memory : ~1000X difference in resistivity**

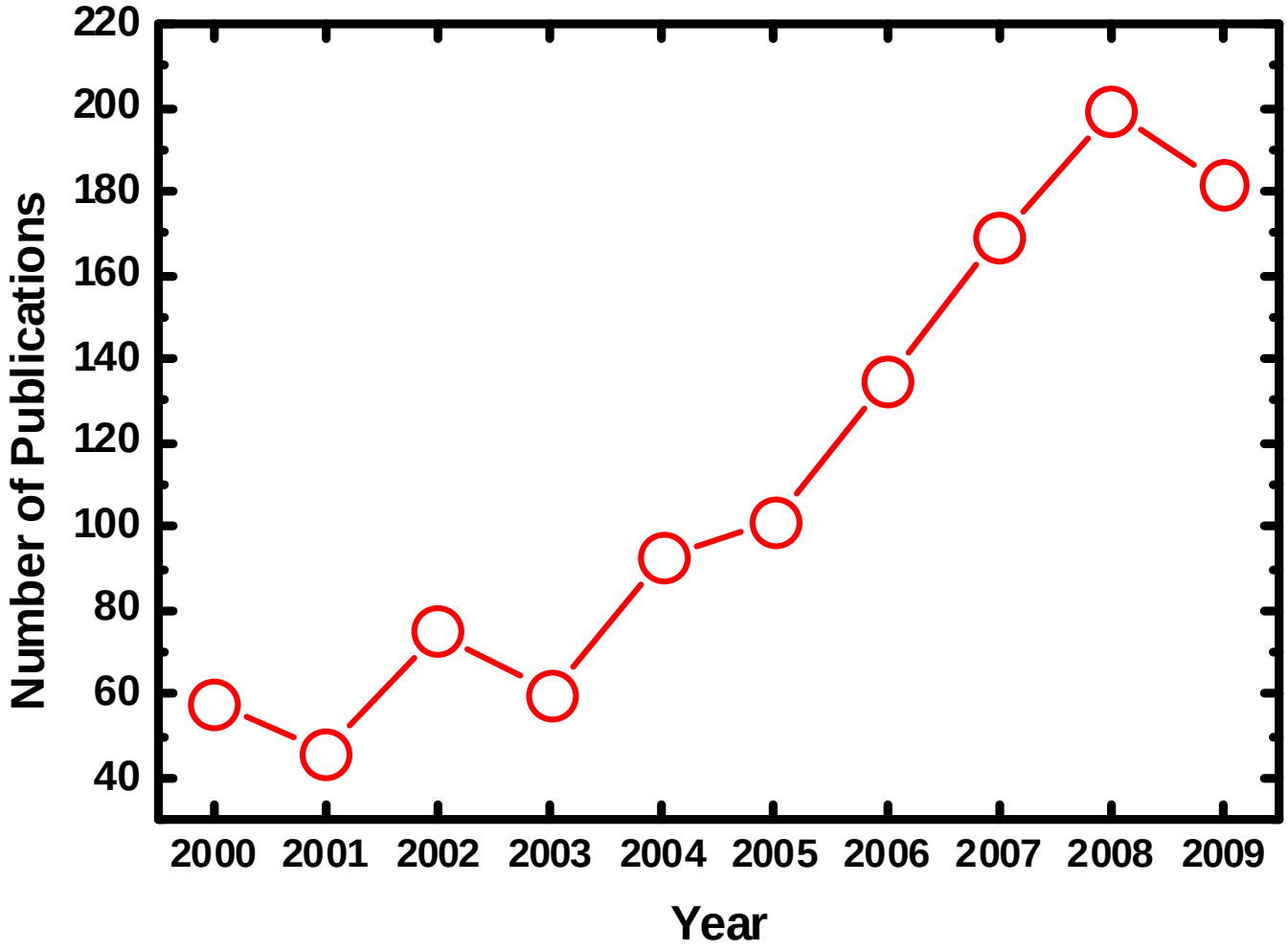
PCM Programming – Joule heating





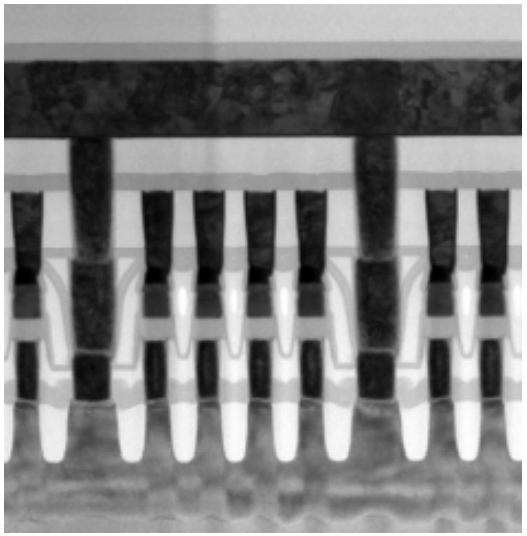
- **Amorphization (RESET): Melt and quench ($T>T_{Melt}$)**
- **Crystallization (SET): Anneal ($T>T_{crys}$)**

Phase Change Memory

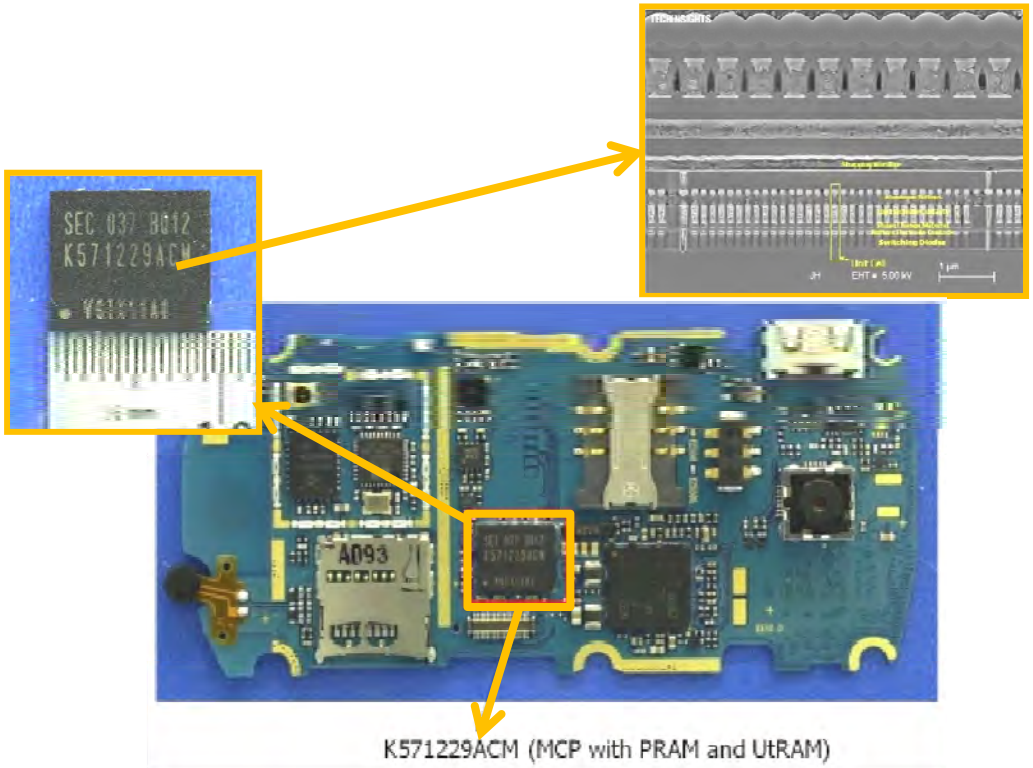


H.-S. P. Wong et al., “Phase Change Memory,” *Proc. IEEE* 2010.

PCM Status

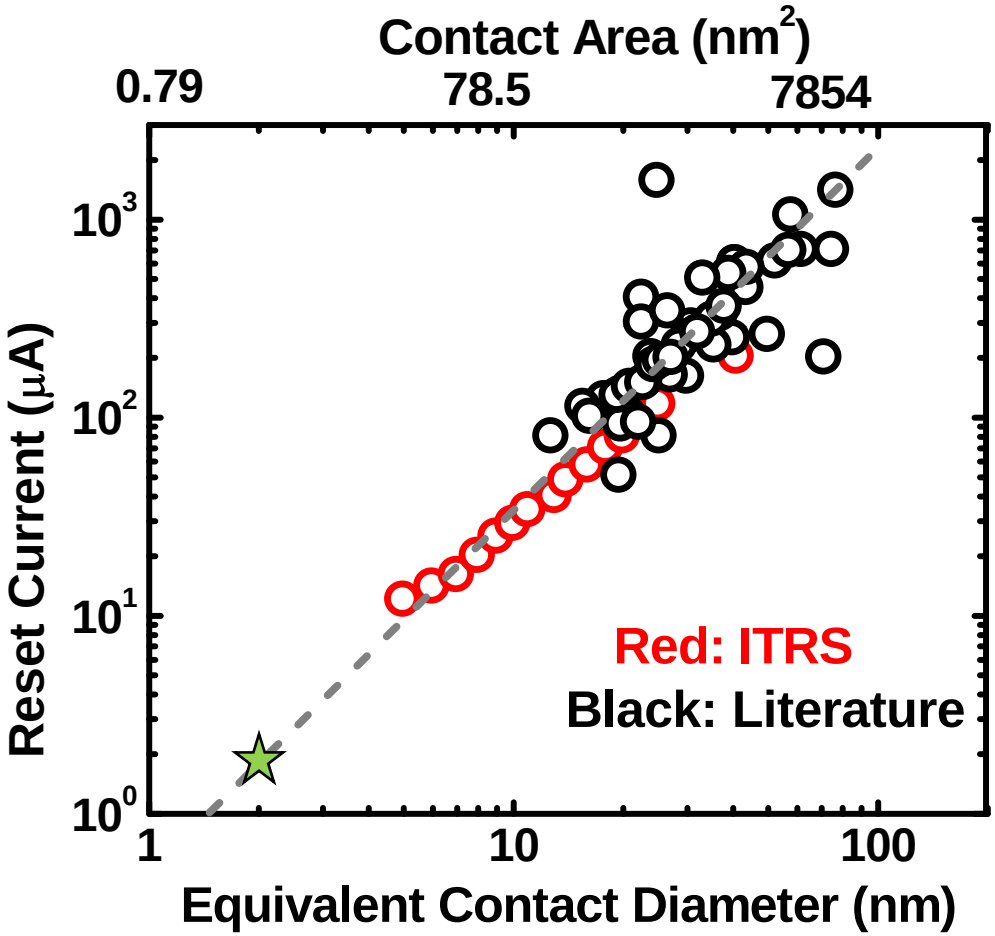


Numonxy 1 Gb PCM array
45 nm generation
(ISSCC, 2010)



512-Mbit PRAM die packaged with a Samsung 128Mbit UtRAM die in a multi-chip package in a Samsung mobile phone (Dec, 2010)

Scaling of Programming Current



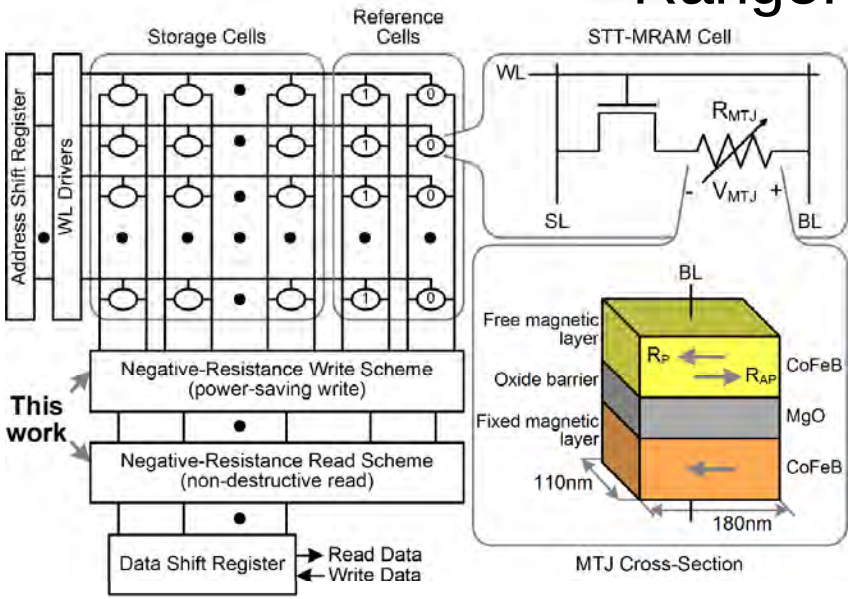
Current density = 20 – 40 MA/cm^2

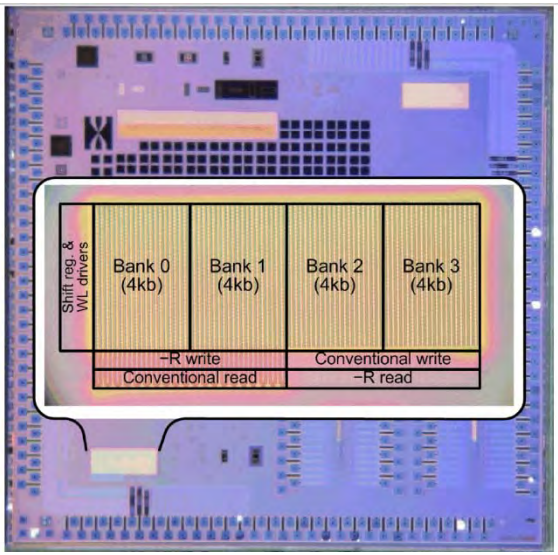
(can be 10X lower with materials engineering)

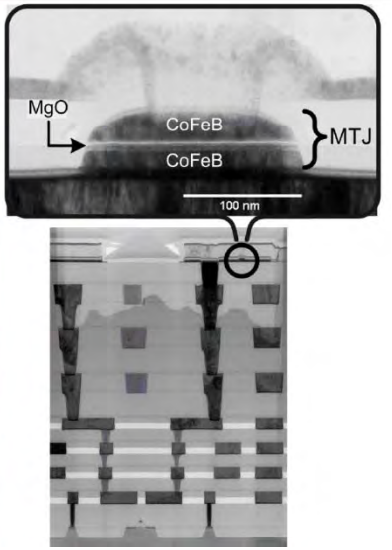
J. Liang, R. G.D. Jeyasingh, H.-Y. Chen, H.-S. P. Wong, "A 1.4μA Reset Current Phase Change Memory Cell with Integrated Carbon Nanotube Electrodes for Cross-Point Memory Application," Symp. VLSI Technology, paper 5B-4, Kyoto, Japan, June 13 – 16, 2011.

STTRAM

Write current density = 2MA/cm^2
Range: 3 – 5 MA/cm^2



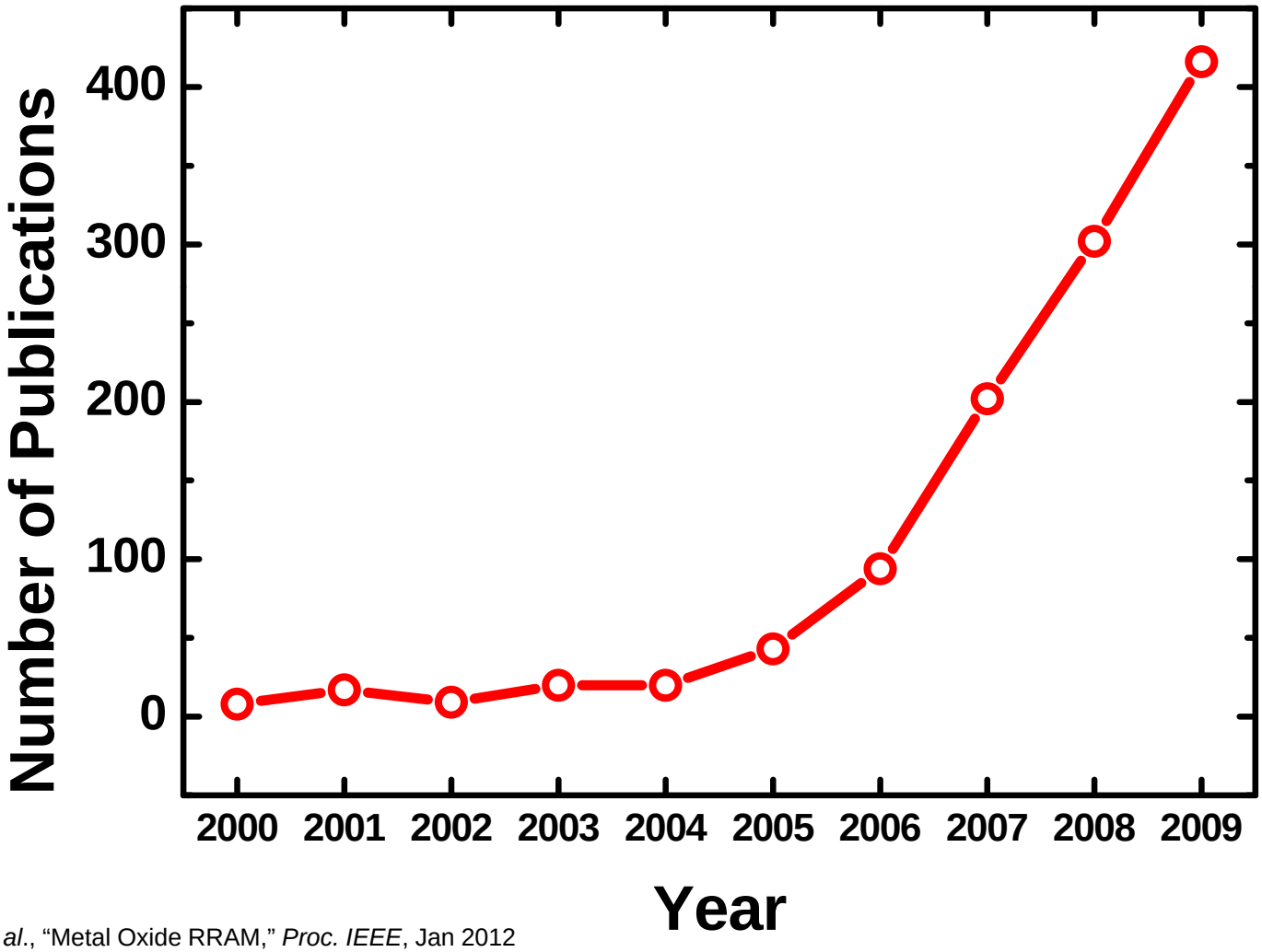




		ISSCC 2007 [1]	This Work
Process		0.2μm CMOS (1P, 4M)	0.13μm CMOS (1P, 8M)
Read time		40ns	**8ns**
Write	**time**	100ns	**9/10ns**
	current	0.2mA	0.40 / 0.87mA
Capacity		2Mb	16kb
Memory Cell Sizes		1.6 x 1.6μm^2	3.25 x 1.7μm^2
Power Supply		1.8V	1.2V:read and logic/3.3V:memory core

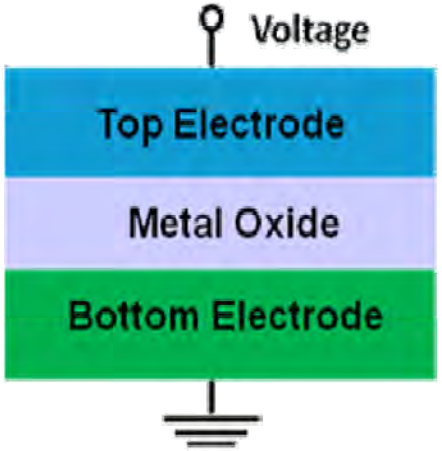
D. Halupka et al., "Negative-Resistance Read and Write Schemes for STT-MRAM in 0.13μm CMOS," ISSCC, p. 256 (2010). (U. Toronto/Fujitsu)

RRAM is Hot!



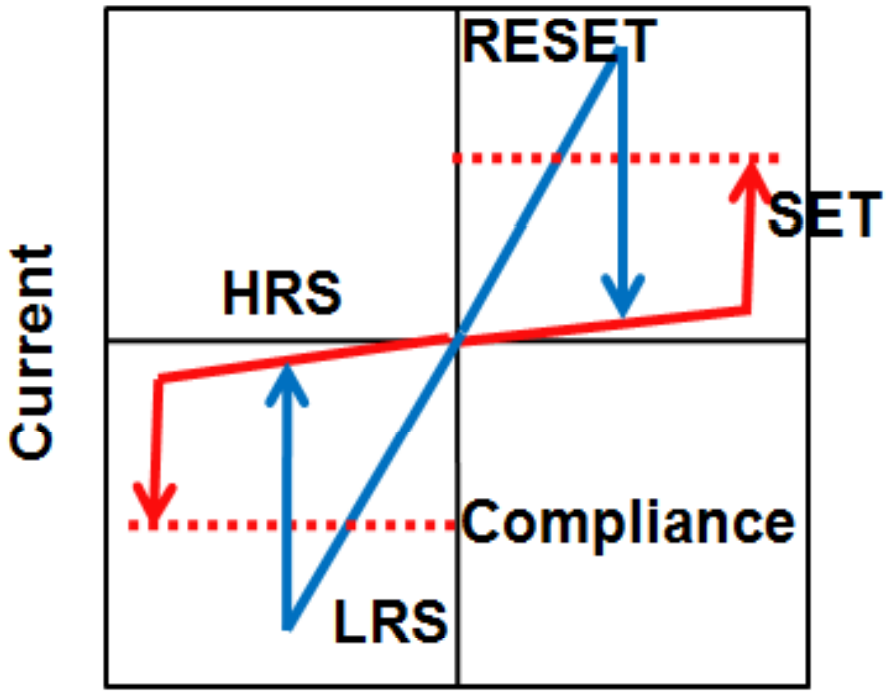
H.-S. P. Wong *et al*., “Metal Oxide RRAM,” *Proc. IEEE*, Jan 2012

Metal Oxide M-I-M Memory (RRAM)

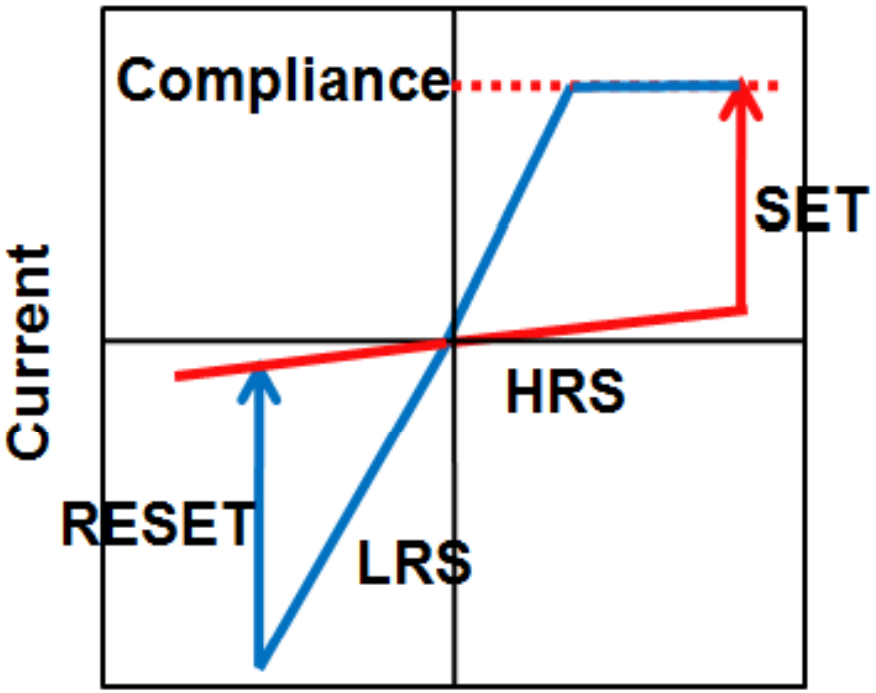


- **Motivation:**
  - Low programming voltage (< 3V)
  - Material set compatible with conventional semiconductor processing (e.g Ni, Hf, Al…)
  - Low temperature processing (BEOL-compatible)
- **Key issues:**
  - Physics of resistive switching
  - Device scaling properties
  - Device uniformity
  - A killer application

Basic I-V Characteristics

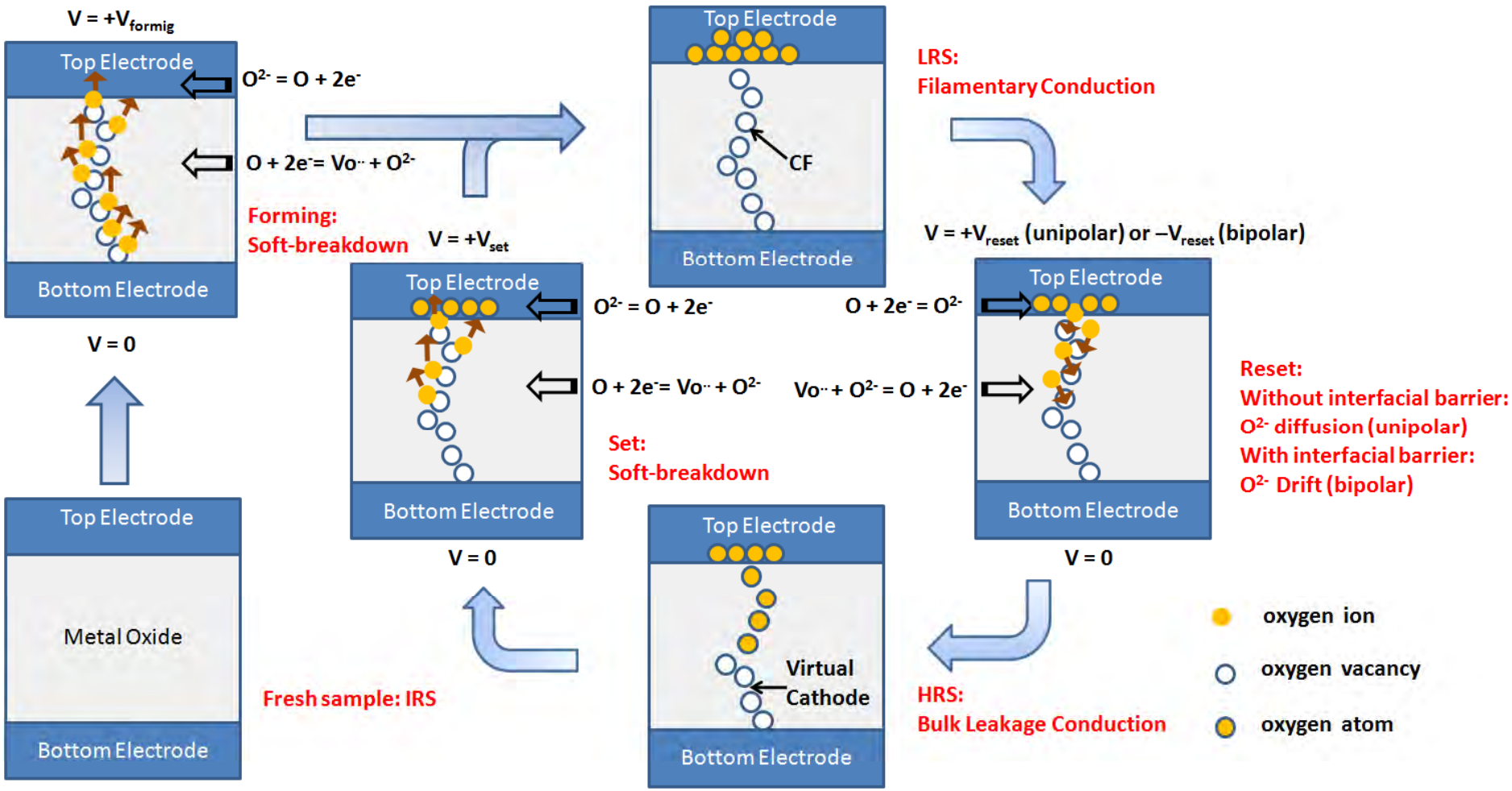


unipolar mode



bipolar mode

Prevailing Theory for Resistive Switching



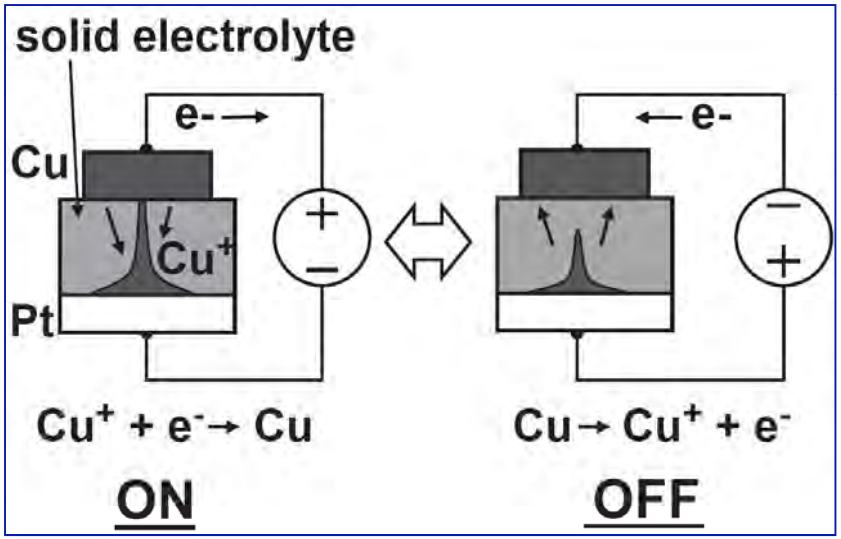
S. Yu, B. Lee, H.-S. P. Wong, "Metal Oxide Memory," in J. Wu, W. Han, H.-C. Kim, A. Janotti eds, "Functional Metal Oxide Nanostructures," Springer 2011.

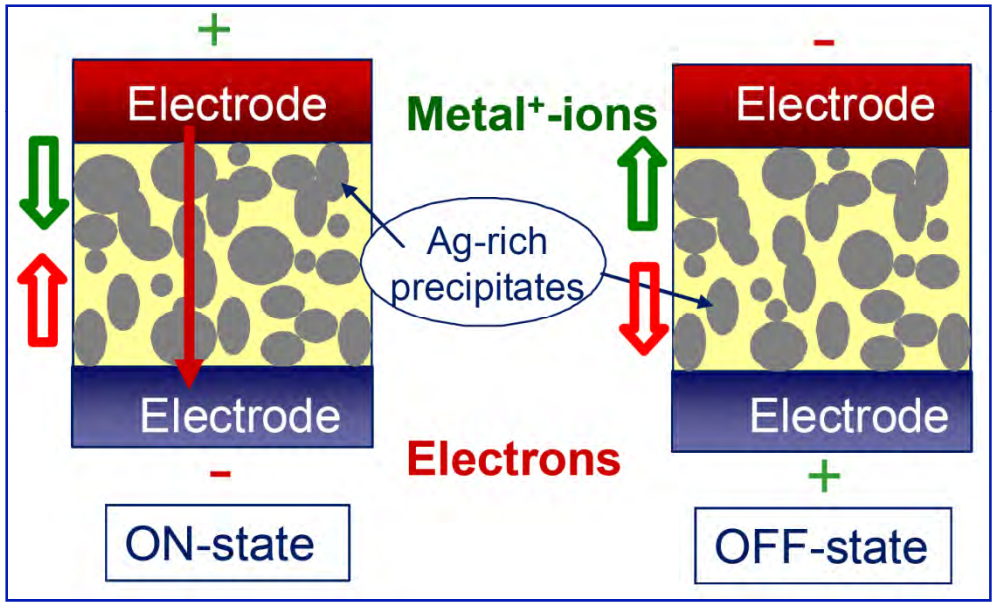
Conductive Bridge Memory

1. Redox reaction
2. Ion migration (cation toward cathode)

$M \rightarrow M^+ + e^-$

$M \leftarrow M^+ + e^-$



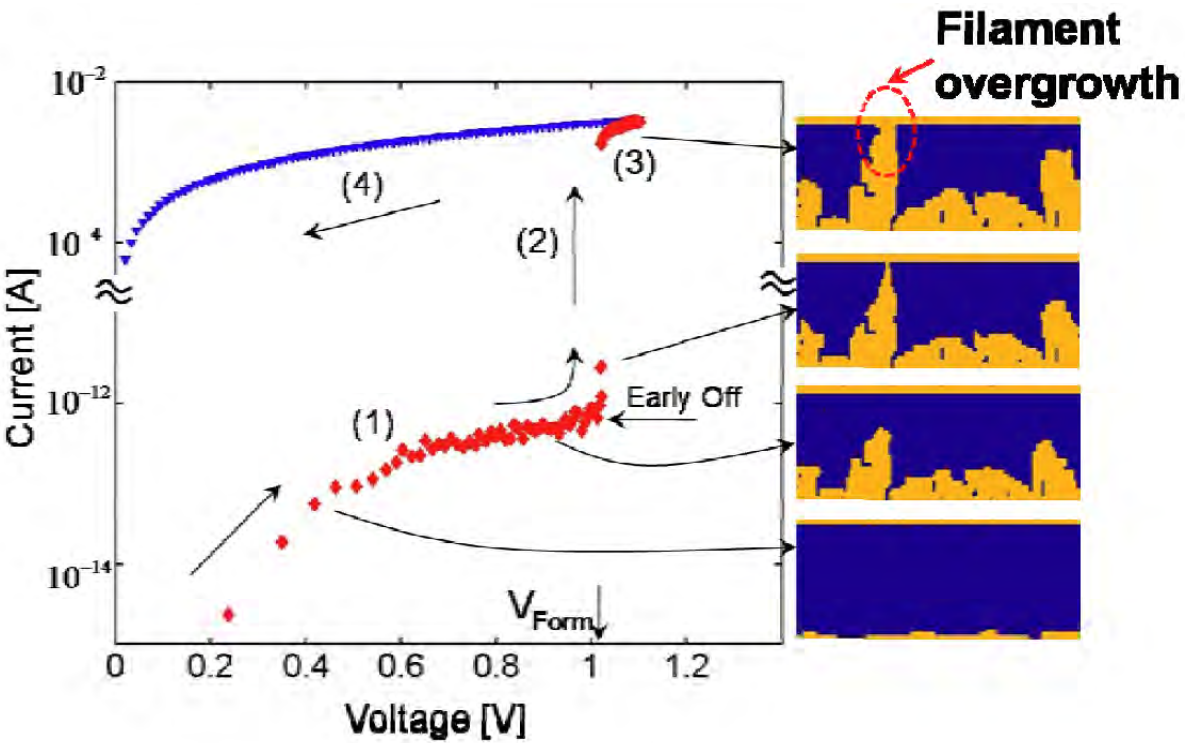


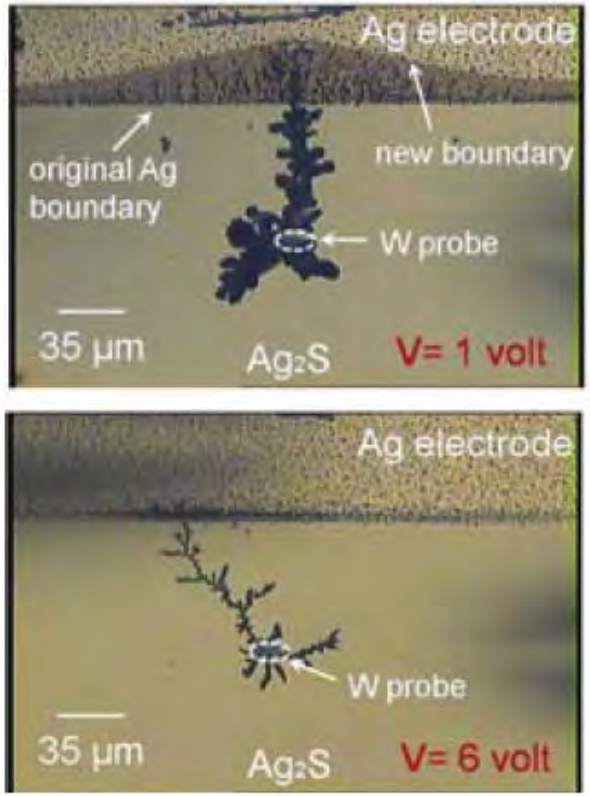
Cu/Cu_2S

Ag-Ge-Se

M. Kund et al., "Conductive bridging RAM (CBRAM): an emerging non-volatile memory technology scalable to sub 20 nm," IEDM, p. 754 (2005).
T. Sakamoto et al., "A Ta_2O_5 solid-electrolyte switch with improved reliability," Symp. VLSI Technology, p. 38 (2007).

CBRAM Filament Formation



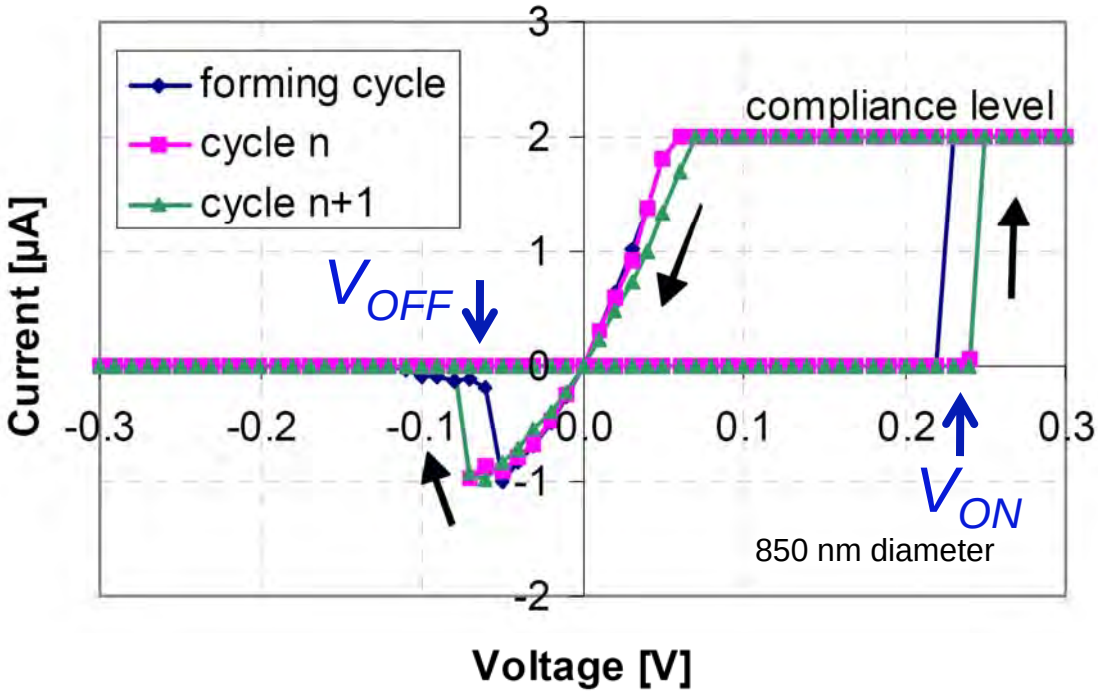


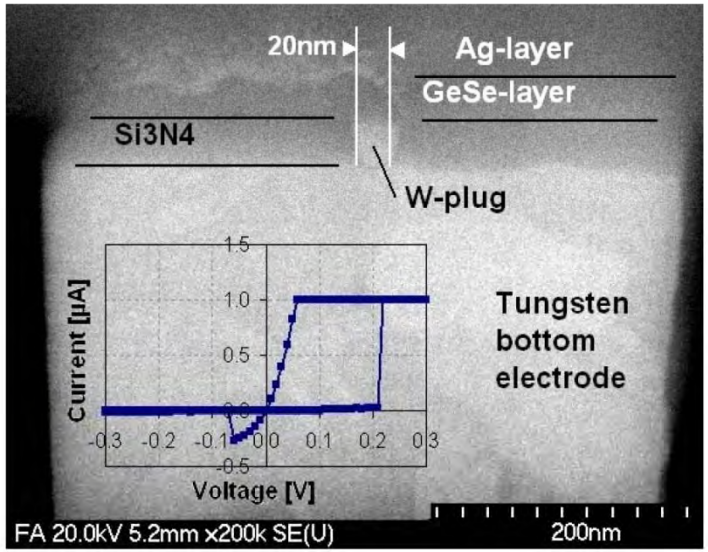
F. Pan, S. Yin, V. Subramanian, "A comprehensive simulation study on metal conducting filament formation in resistive switching memories," p. 111, *IMW* 2011.

Nanoconductive Bridge Typical I-V Characteristics

- **Bipolar, asymmetric programming/erase**
- **Low programming/erase voltage**
- **Forming required**

• Ag top electrode (active electrode)
• GeSe chalcogenide glass as solid electrolyte
• W bottom electrode (inert electrode)



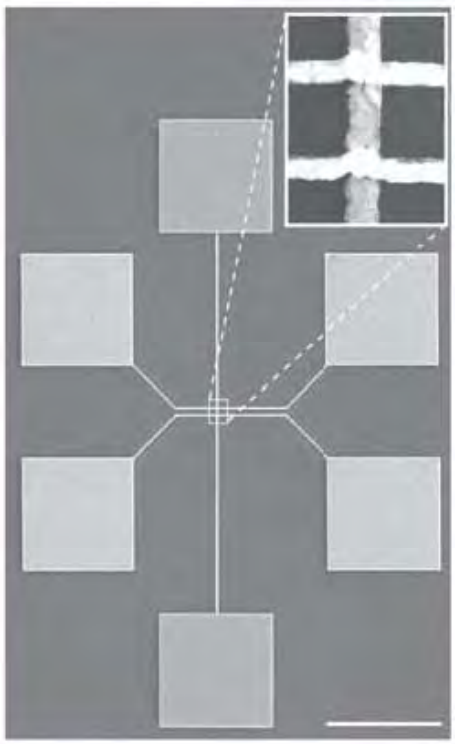


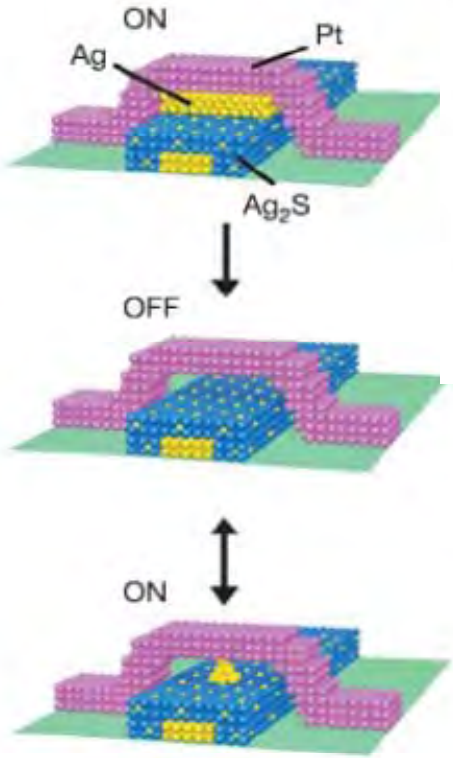
M. Kund et al., "Conductive bridging RAM (CBRAM): an emerging non-volatile memory technology scalable to sub 20 nm," IEDM, p. 754 (2005).

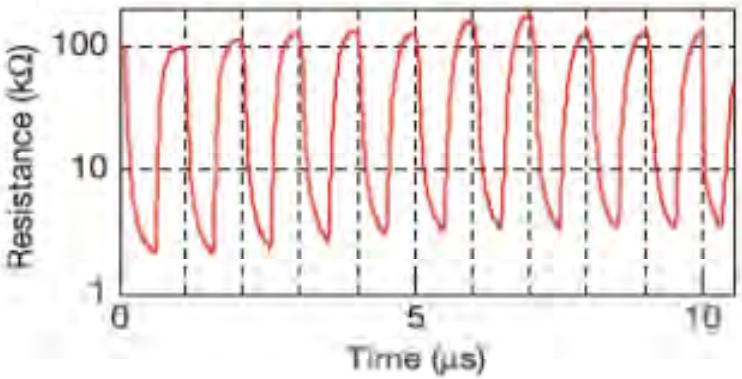
$Ag/Ag_2S/Pt$

$$Ag^+(Ag_2S) + e^- \underset{\text{Oxidation}}{\overset{\text{Reduction}}{\rightleftharpoons}} Ag$$

An oxidizable electrode (Ag, Cu *etc.*) is needed to fulfill the forming and dissolving filamentary conduction paths







Switching at (±0.6 V,1 MHz)

Top: As-formed switched-on state

Middle: Switched-off state

Bottom: Switched-on state after the initial switching-off process (bottom).

K. Terabe *et al.*, "Quantized conductance atomic switch", *Nature* 433, 47-50 (2005).

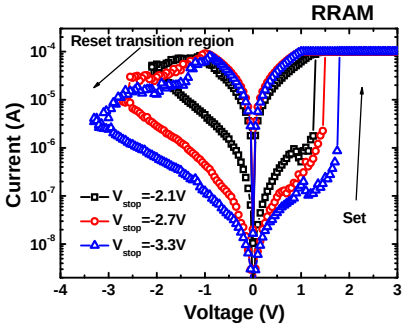
So, Let's Talk About Energy

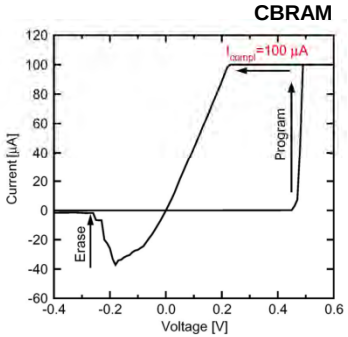
- **Reading a bit**
  - Determined by the resistance of the lowest resistance state (LRS)
  - Typical: read voltage=200mV, read time = 10 ns
  - LRS: RRAM = 10kΩ – 100MΩ, CBRAM = kΩ – GΩ, STTRAM = few kΩ, PCM = kΩ – MΩ
  - Read energy ~ 0.4pJ – 4aJ
- **Writing a bit**
  - Depends on physics and materials

Write Energy – PCM and STTRAM

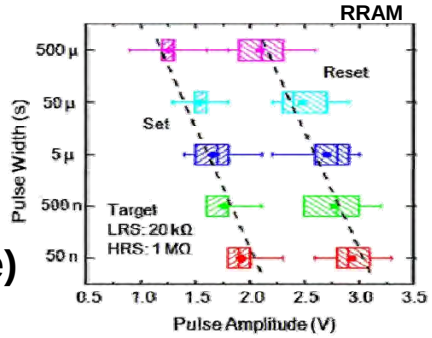
- **Write voltage roughly constant**
  - PCM ~ 1 – 2V
  - STTRAM ~ 0.5 – 1V
- **Write current density**
  - PCM: 20 – 40 MA/cm^2
  - STTRAM: 3 – 5 MA/cm^2
- **Write speed**
  - PCM: 10 – 30 ns
  - STTRAM: 1 – 10 ns
- **Write energy per bit (currently, scales with device size)**
  - PCM: 3 pJ/bit (0.3 pJ/bit experimentally demonstrated)
  - STTRAM: ~0.3 pJ/bit

Write Energy – RRAM and CBRAM



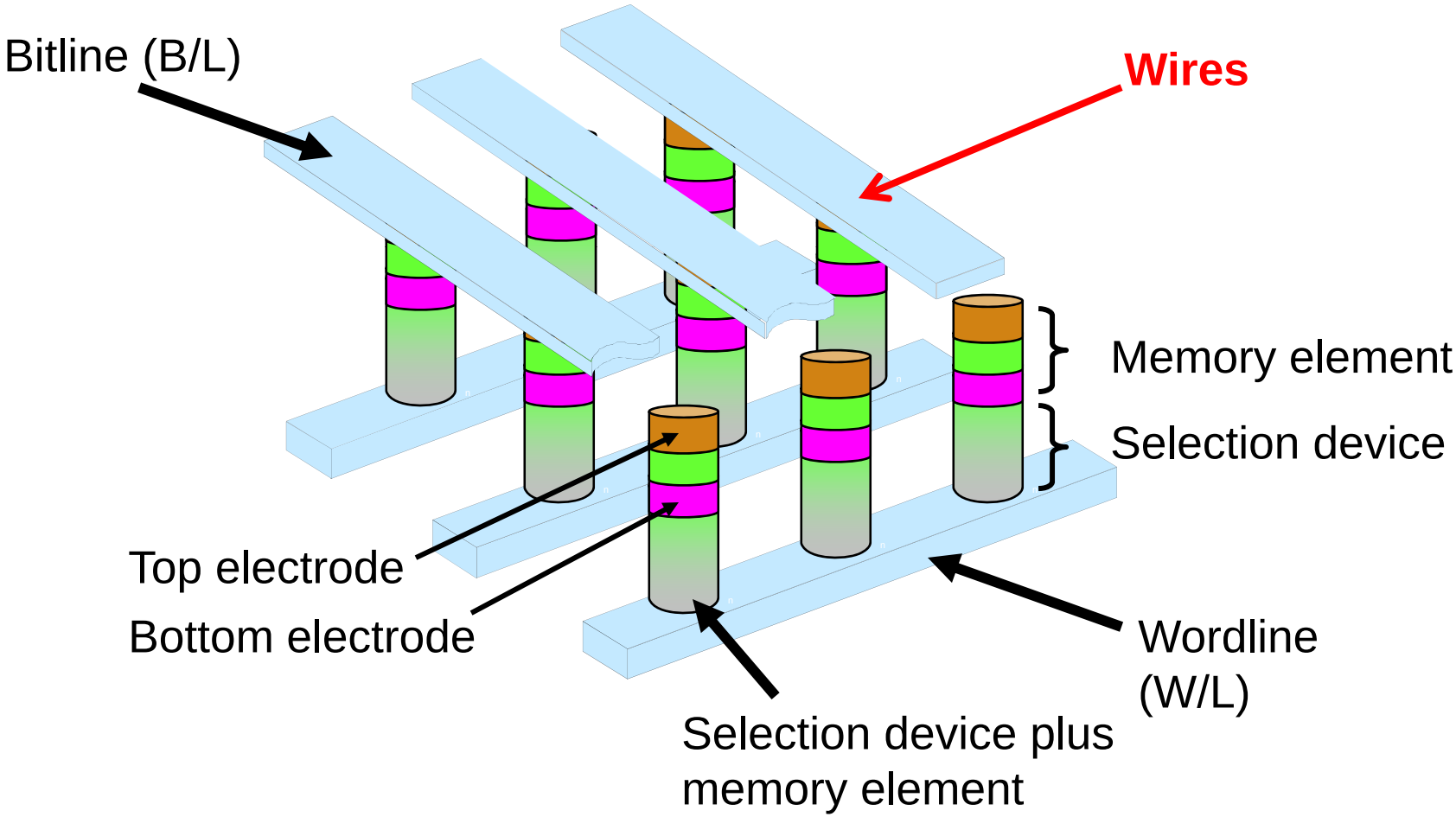


- **Write voltage roughly constant**
  - RRAM ~ 1.5 – 2V ($V_{RESET} > V_{SET}$)
  - CBRAM ~ 1.5 – 3V ($V_{SET} > V_{RESET}$)
- **Write current**
  - Almost a free variable, depends on the resistance value of the low resistance state (LRS)
  - LRS can be "freely adjusted" by controlling the programming current for the LRS
  - LRS: RRAM = 10kΩ – 100MΩ, CBRAM = kΩ – GΩ
  - RRAM ~ µA (or even 10 nA) , CBRAM ~ 100 µA
  - Tradeoff between lower current (energy) and read current (time) and retention
- **Write speed**
  - Write time decreases exponentially with write voltage
  - Energy = $(V^2/R_{LRS}) \times t$
  - So, use large voltage amplitude to decrease time to reduce energy
- **Write energy per bit (filamentary, does NOT scale with device size)**
  - RRAM: 60 fJ/bit – 3 pJ/bit
  - CBRAM: 1 – 5 pJ/bit



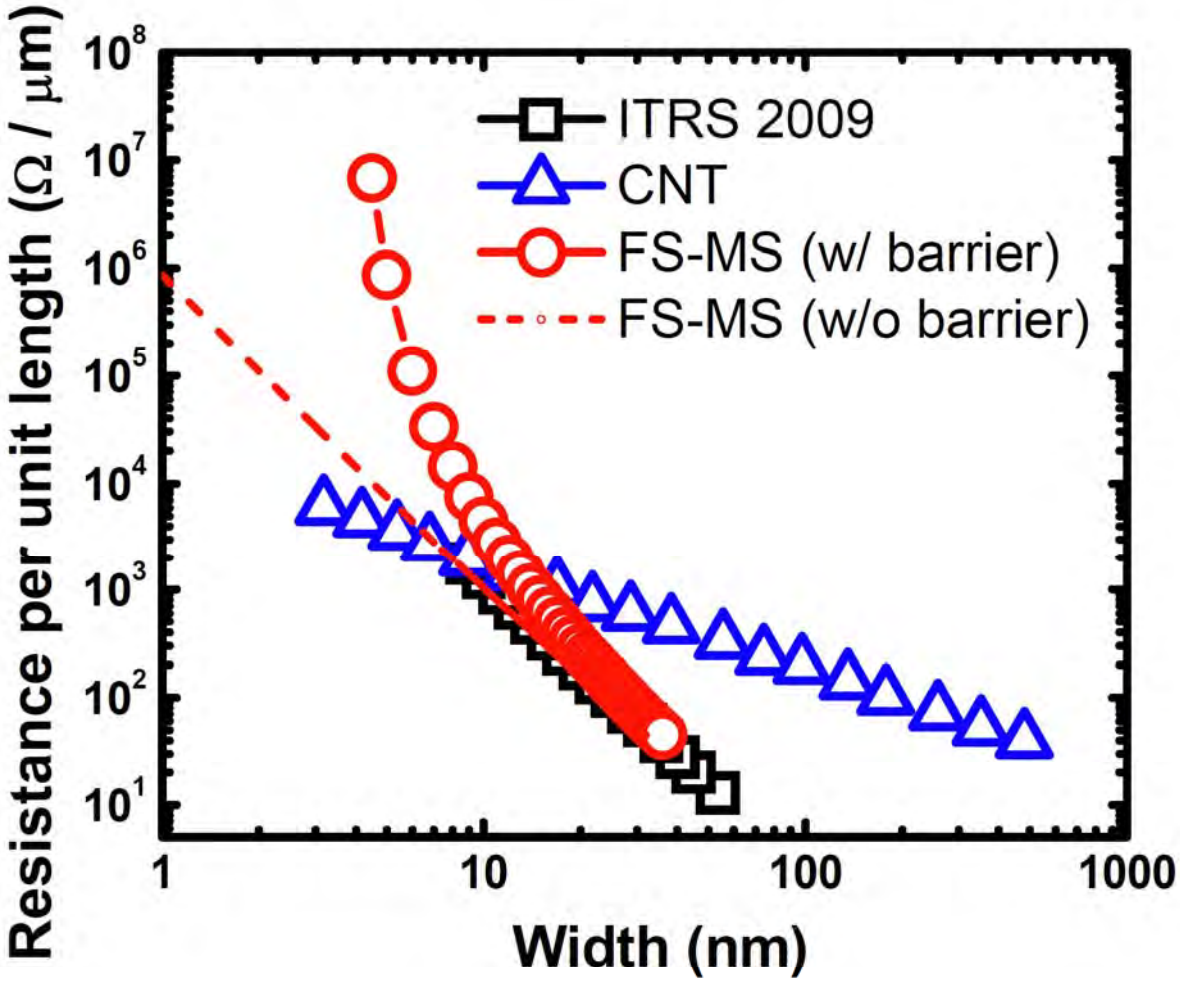
RRAM: S. Yu. Y. Wu, H.-S. P. Wong, APL, 103514, 2011
CBRAM: U. Russo, D. Kamalanathan, D. Ielmini, A. Lacaita, M. Kozicki, EDL, p. 1040, 2009.

Cross-Point Memory Array with Selection Device



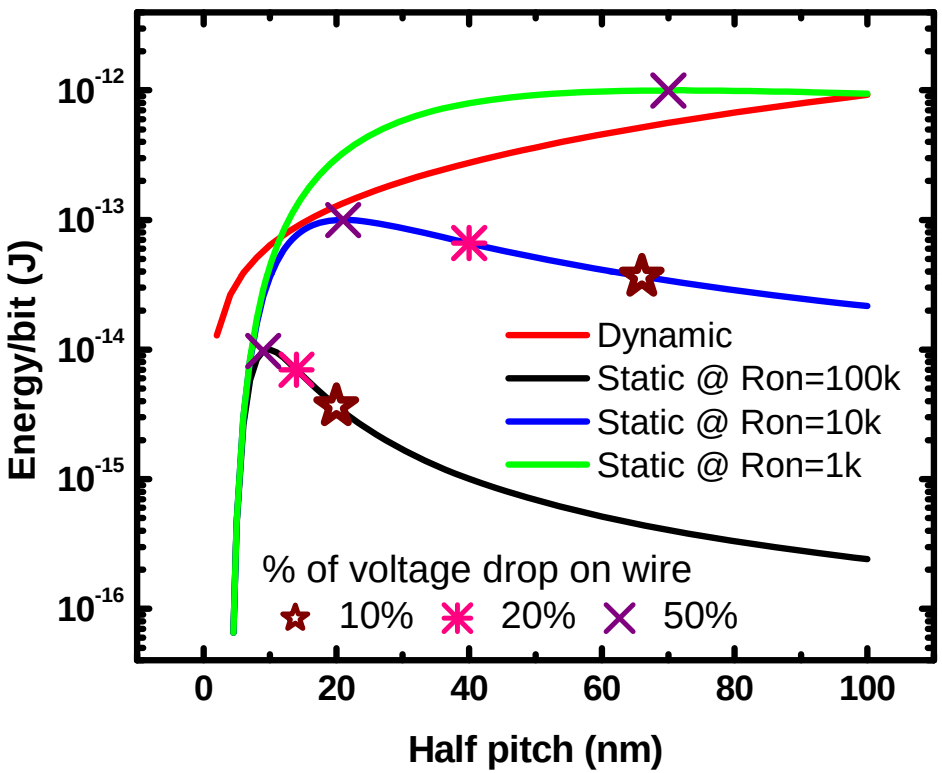
Let’s talk about wires

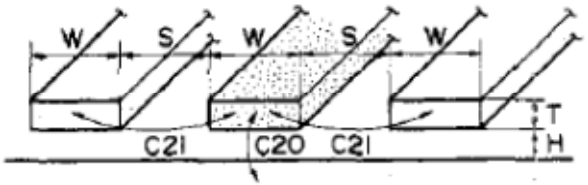
Bitline/Wordline Resistance



J. Liang, Y. Wu, H.-S. P. Wong, *ACM JETC* (submitted)

Energy Consumed in Wires





$$C_{tot} = \frac{l}{2} \times C_{20} + l \times 2C_{21} \quad (F)$$

Charging wires = CV^2

Static energy = $I \times V \times t$

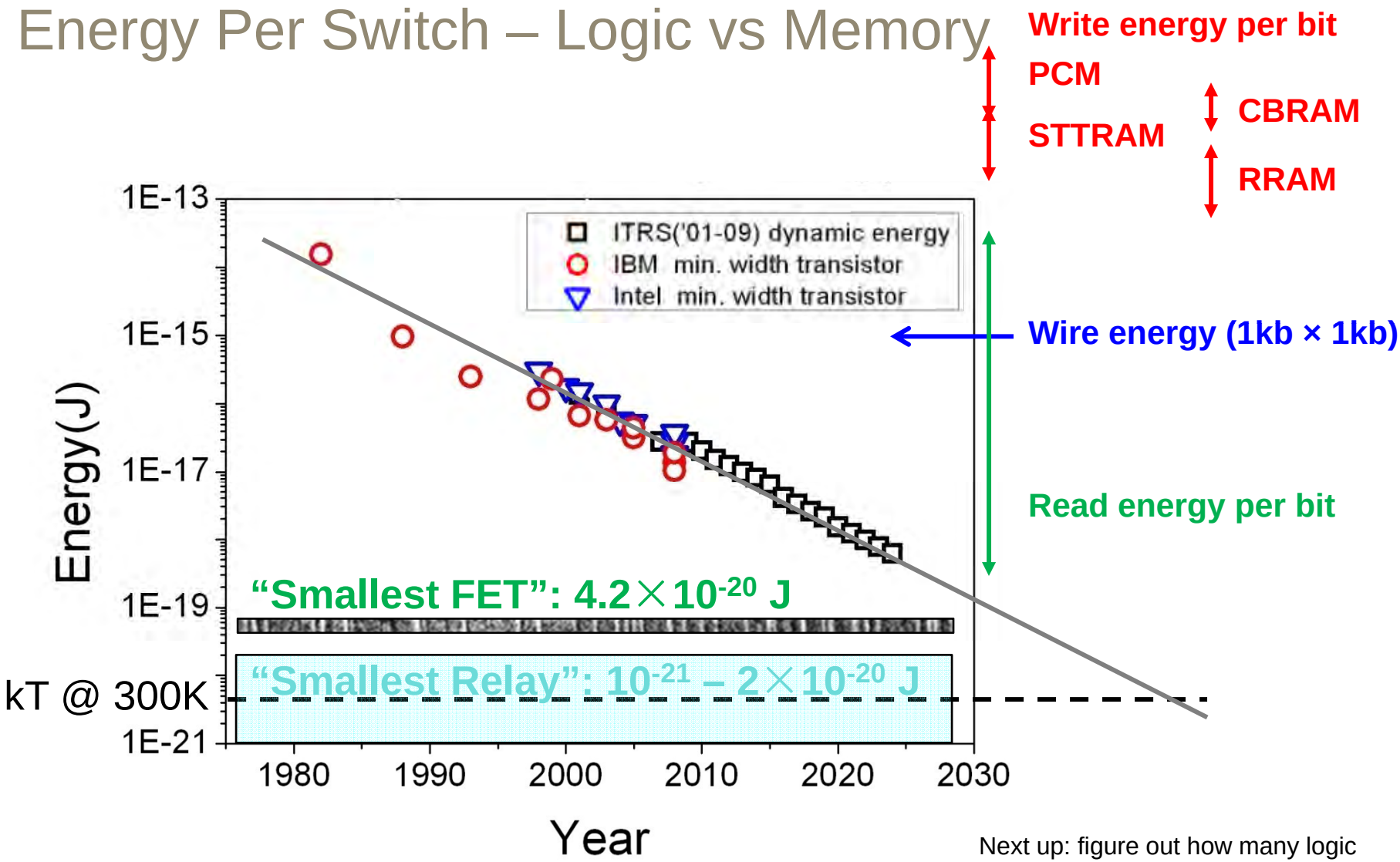
$$= \frac{V}{R_{wire} + R_{on}} \times \frac{R_{wire} \times V}{R_{wire} + R_{on}} \times t$$

Charging Energy ≈ Static Energy

Energy consumed in wires ~ 100X smaller than devices (not including readout circuits)

J. Liang (2011)

Energy Per Switch – Logic vs Memory



Next up: figure out how many logic transitions per memory write

Students and Post-Docs

NEM Relay

Simon Guan

NEM Relay

Daesung Lee

NEM Relay

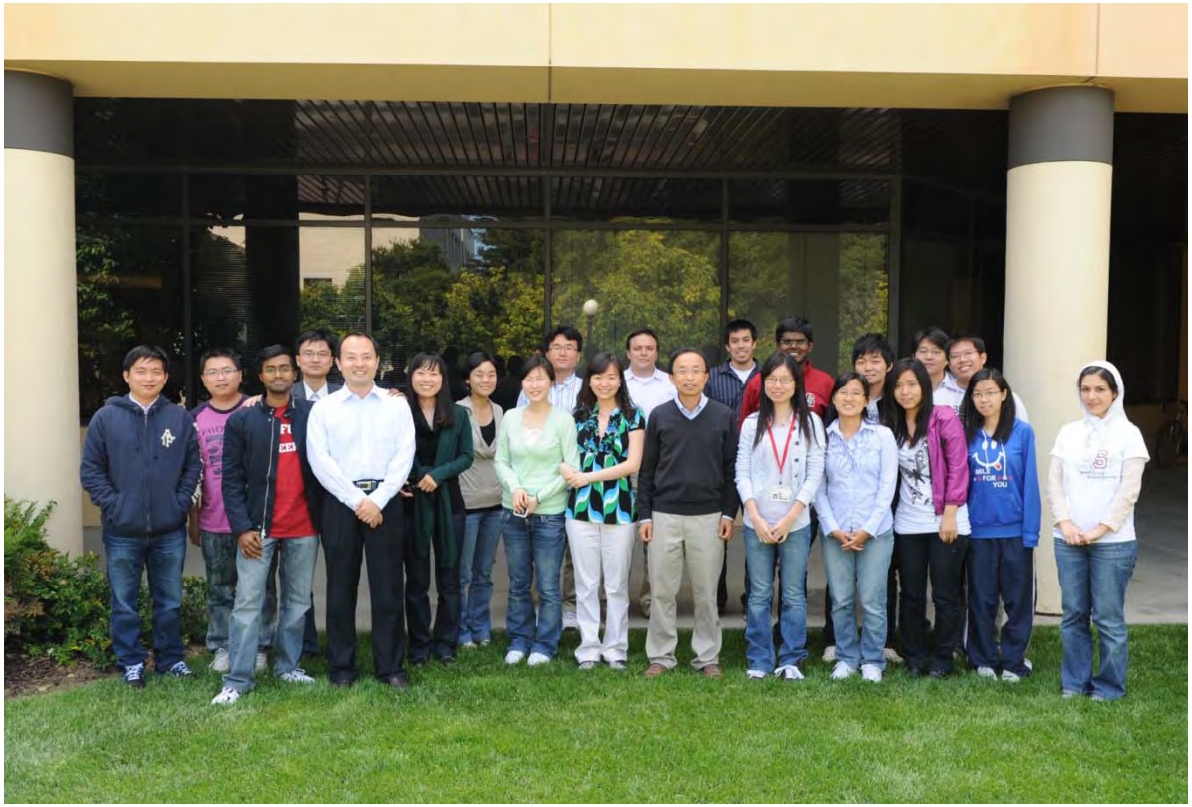
J Provine

Nanotube device

Lucky Liyanagi

Nanotube device

Henry Chen



III-V, Ge, and CMOS

Crystal Kenney

III-V, Ge, and CMOS

Jenny Hu

Nanotube / graphene interconnect

Helen Chen

NEM relay

Xiaoying Shen

Self-assembly

He Yi

Carbon Nanotube

Jason Parker

Memory

Jiale Liang

Phase Change memory

Rakesh Jeyasingh

CMOS device/circuit

Jieying (Ivy) Luo

Chip in Cell

Kokab Parizi

NEM relay

Soogine Chong

Self-assembly

Marissa Caldwell

Phase change memory

Ethan Ahn

Metal oxide memory

Shimeng Yu

Metal oxide memory

Byoungil Lee

Metal oxide memory

Yi Wu

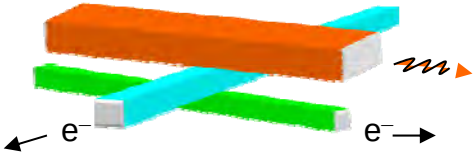
Sponsors and Collaborators

Stanford INMP

(Toshiba, Intel, TI, IBM, AMD, TEL, AMAT, COSAR, Synopsis)

Stanford NMTRI

(Toshiba, Intel, COSAR (Samsung, Hynix), Micron, SanDisk, Intermolecular)



Non-Volatile Memory Technology Research Initiative (NMTRI) at Stanford University

SanDisk
Micron
TOSHIBA
COSAR
SAMSUNG
intel
hynix
INTERMOLECULAR
FAST FORWARD THINKING

Technical Collaborators