EVOLUTION OF
NONVOLATILE SEMICONDUCTOR MEMORY
From Invention to Nanocrystal Memory

S. M. SZE

National Chiao Tung University
National Nano Device Laboratories
Hsinchu, Taiwan

OUTLINE

- INTRODUCTION
- DEVICE PHYSICS
- HISTORICAL DEVELOPMENT OF NVSM
- EMERGING TECHNOLOGIES
- APPLICATIONS AND FUTURE TRENDS
- CONCLUSION

LEMMA OF NEW TECHNOLOGY

“ The principal applications of any sufficiently new and innovative technology always have been — and will continue to be — applications *created* by that technology ” — Herbert Kroemer

EXAMPLES

- **The Transistor (1947)**

  was not just a replacement for vacuum tubes, it *created* the modern computer and the new industrial revolution.

- **The Heterojunction Laser (1963)**

  has revolutionized the optoelectronics technology, it *created* optical fiber communication, CD, and DVD.

- **The Nonvolatile Semiconductor Memory (1967)**

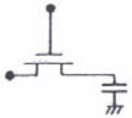
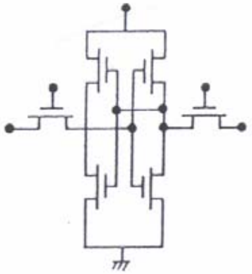
  has revolutionized the information storage technology, it *created* the mobile phone, notebook computer, digital photography, MP3, PDA and GPS.

SEMICONDUCTOR MEMORIES

- **VOLATILE MEMORIES**– Lose stored information once supply is switched off
  - DRAM (Dynamic Random Access Memory)
  - SRAM (Static Random Access Memory)

- **NONVOLATILE MEMORIES** – Keep stored information when the power supply is switched off
  - Flash MEMORY
  - EEPROM (Electrically Erasable-Programmable Read Only Memory)
  - EPROM (Erasable-Programmable Read Only Memory)

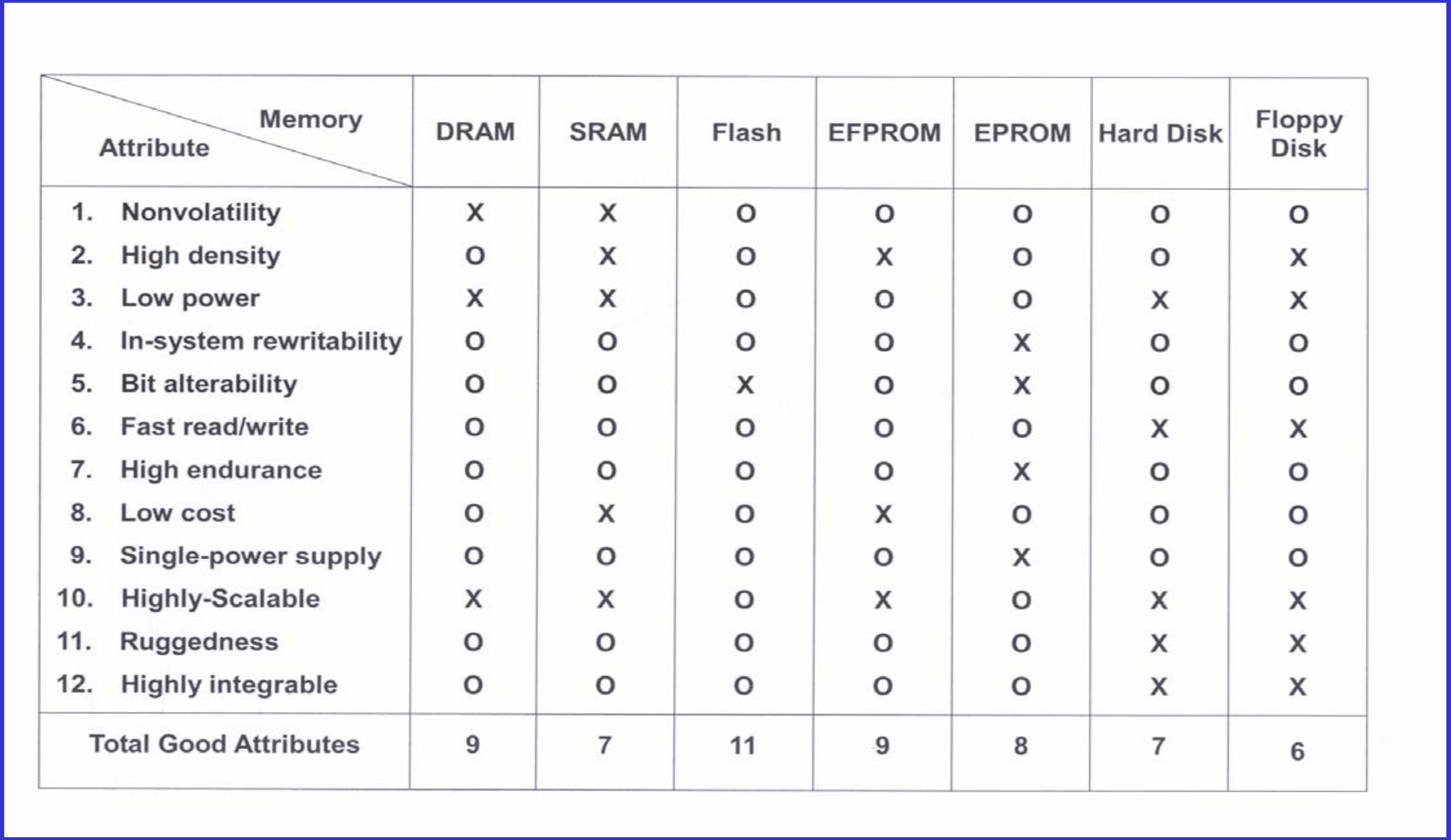
COMPARISON OF SEMICONDUCTOR MEMORIES

DRAM	SRAM	Flash
		
1T + 1C	6T / 4T+2L	1T
Volatile	Volatile	Nonvolatile

ATTRIBUTES OF AN IDEAL MEMORY

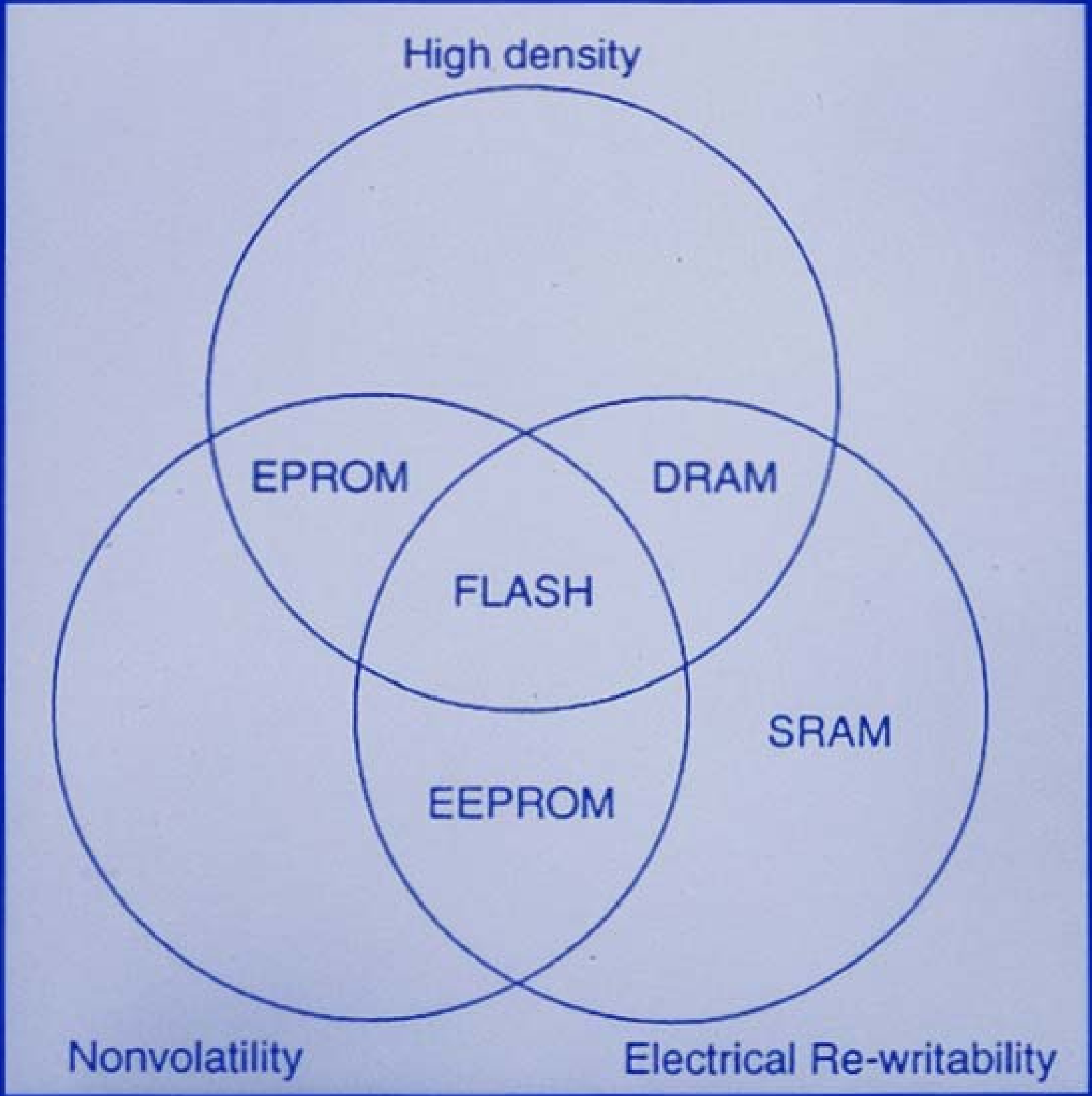
1. Nonvolatility
2. High density (consumes small space/bit)
3. Low power consumption
4. In-system rewritability
5. Bit alterability
6. Fast read/write/erase
7. Endurance (high write/erase cycles)
8. Low cost
9. Single-power supply
10. Highly scalable
11. Ruggedness
12. Highly integrable with other system technologies

COMPARISON OF MEMORY ATTRIBUTES



Memory Attribute	DRAM	SRAM	Flash	EFPROM	EPROM	Hard Disk	Floppy Disk
1. Nonvolatility	X	X	O	O	O	O	O
2. High density	O	X	O	X	O	O	X
3. Low power	X	X	O	O	O	X	X
4. In-system rewritability	O	O	O	O	X	O	O
5. Bit alterability	O	O	X	O	X	O	O
6. Fast read/write	O	O	O	O	O	X	X
7. High endurance	O	O	O	O	X	O	O
8. Low cost	O	X	O	X	O	O	O
9. Single-power supply	O	O	O	O	X	O	O
10. Highly-Scalable	X	X	O	X	O	X	X
11. Ruggedness	O	O	O	O	O	X	X
12. Highly integrable	O	O	O	O	O	X	X
Total Good Attributes	9	7	11	9	8	7	6

COMPARISON OF MEMORY ATTRIBUTES



High density

EPROM

DRAM

FLASH

SRAM

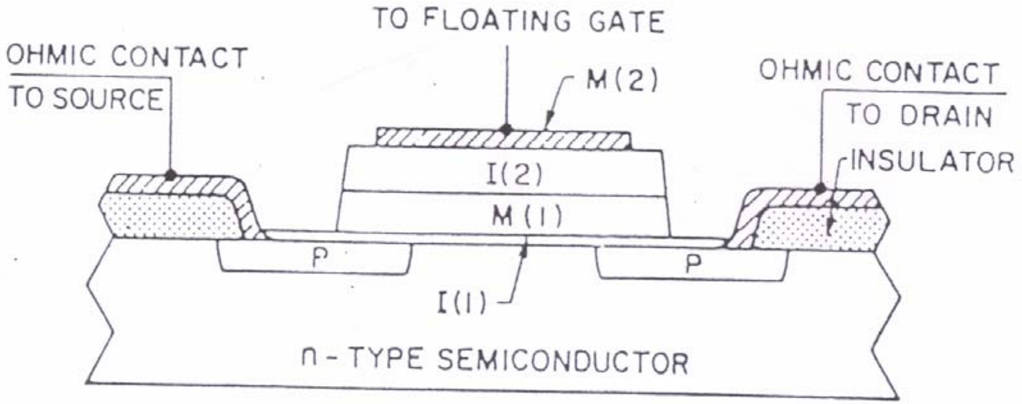
EEPROM

Nonvolatility

Electrical Re-writability

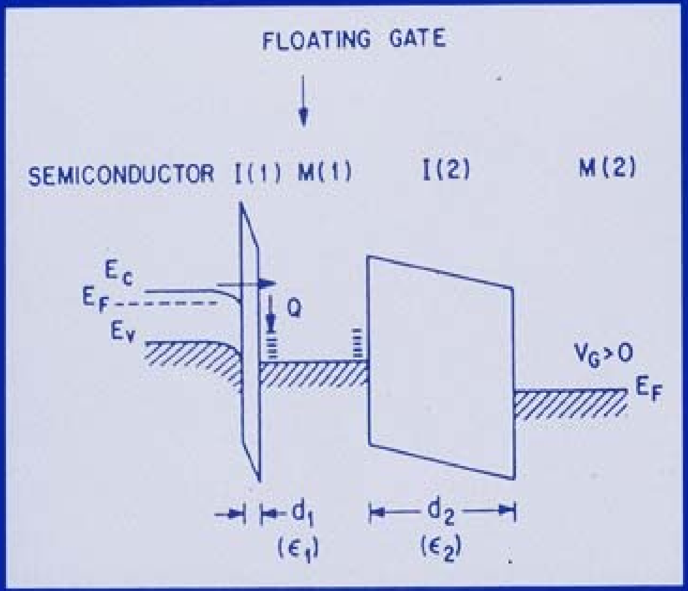
THE FIRST NONVOLATILE SEMICONDUCTOR MEMORY

- The Floating-Gate Concept (1967)

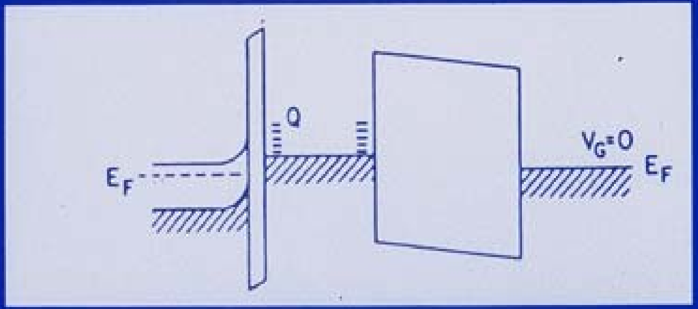


D. Kahng, S. M. Sze "A Floating Gate and its Application to Memory Device" Bell Syst. Tech. J., 46, 1288 (1967)

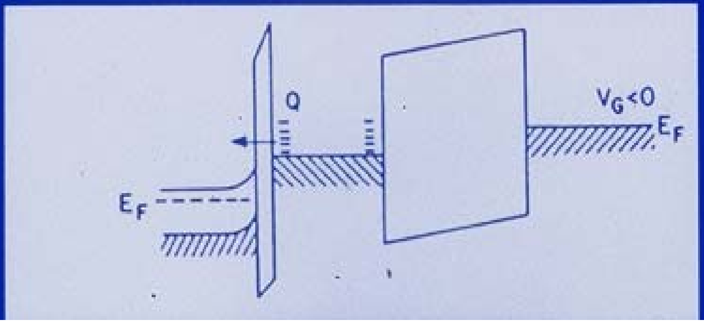
FLOATING-GATE MEMORY DEVICE OPERATION



(a) Programming Mode (Fowler-Nordheim or Direct Tunneling)

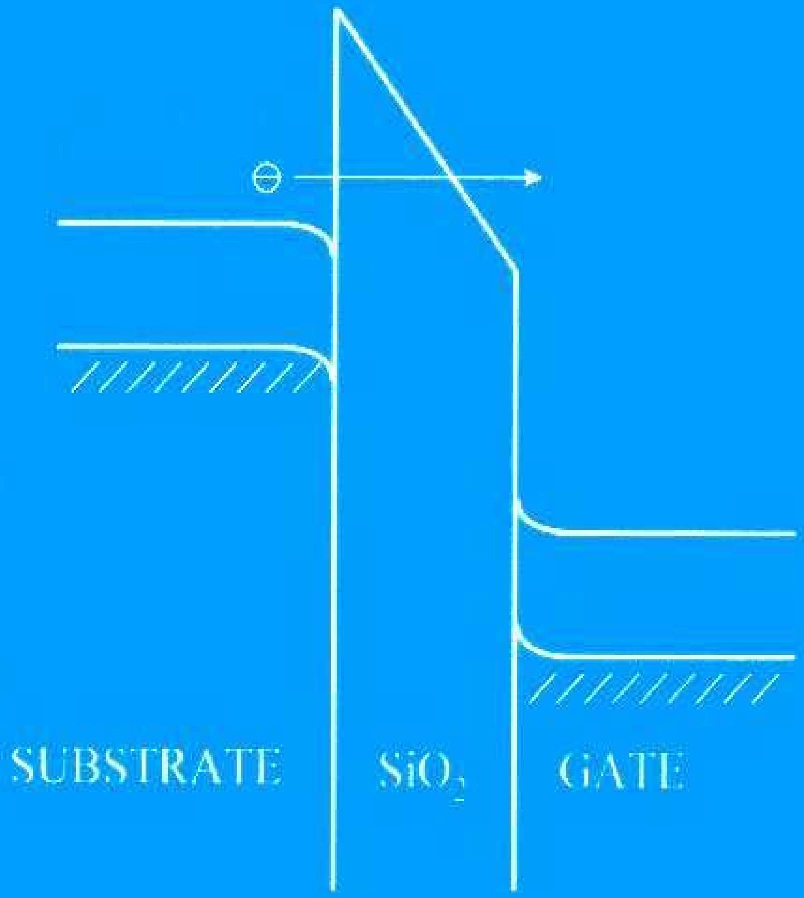


(b) Storage Mode

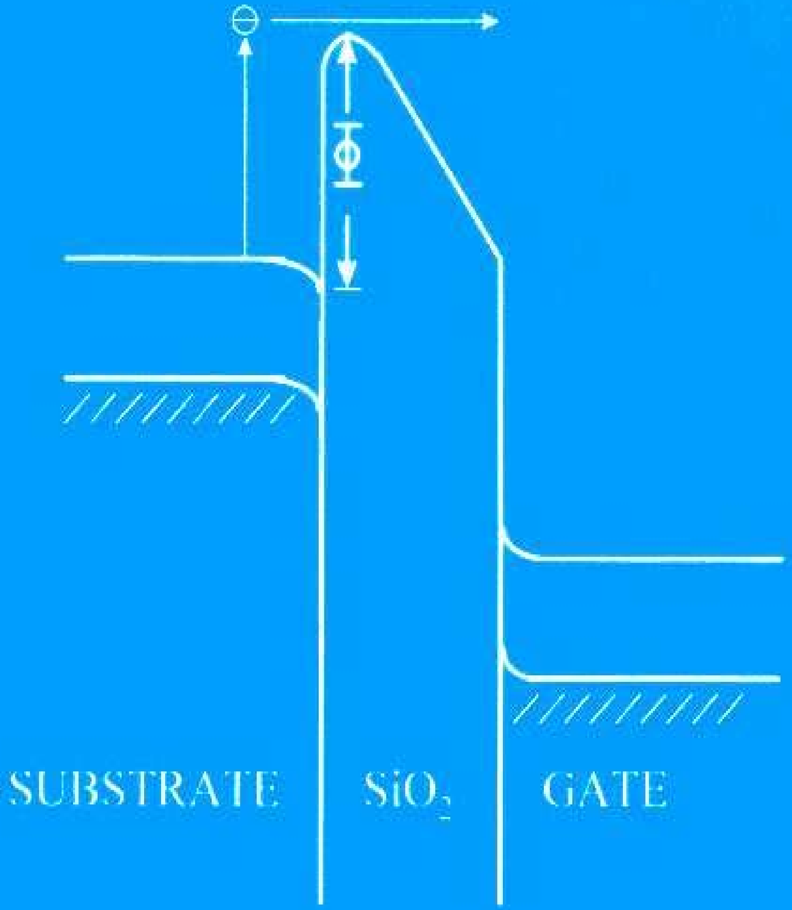


(c) Erase Mode (FN or DT)

CHARGE INJECTION MECHANISMS



(a) FOWLER-NORDHEIM TUNNELING



(b) HOT-ELECTRON INJECTION

CURRENT DENSITY EQUATIONS

For Fowler-Nordheim Tunneling

$$J = C_1 \mathrm{E}^2 \exp\left(-\mathrm{E}_0 / \mathrm{E}\right)$$

$$\mathrm{E} = \frac{V_G}{d_1 + d_2 (\epsilon_1 / \epsilon_2)} - \frac{Q}{\epsilon_1 + \epsilon_2 (d_1 / d_2)}$$

$$Q = \int_0^t J dt' = \text{stored charge}$$

For Hot-Electron Injection

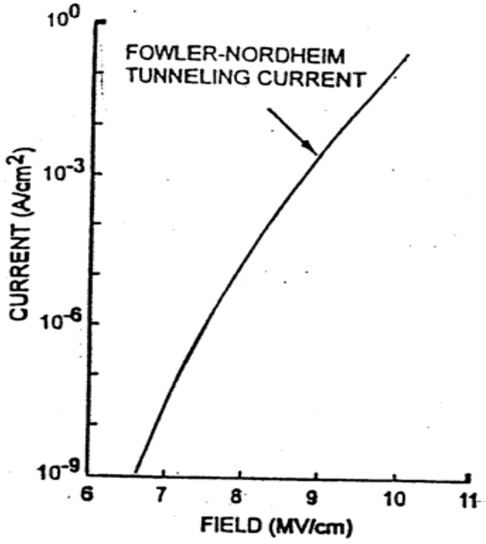
$$J = C_2 \mathrm{E}_m^2 \exp\left(-\Phi / \lambda \mathrm{E}_m\right)$$

E_m = Maximum Channel Lateral Field

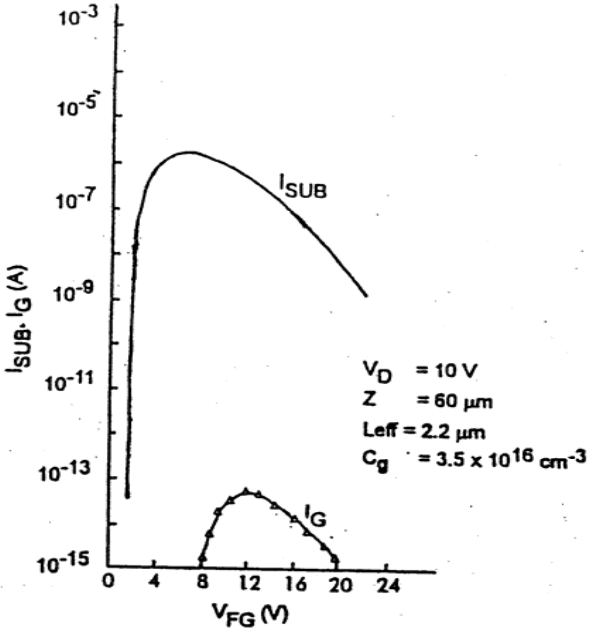
λ = Inelestic Scattering Length

Φ = Barrier Height between Si and SiO_2

GATE CURRENT VERSUS GATE VOLTAGE

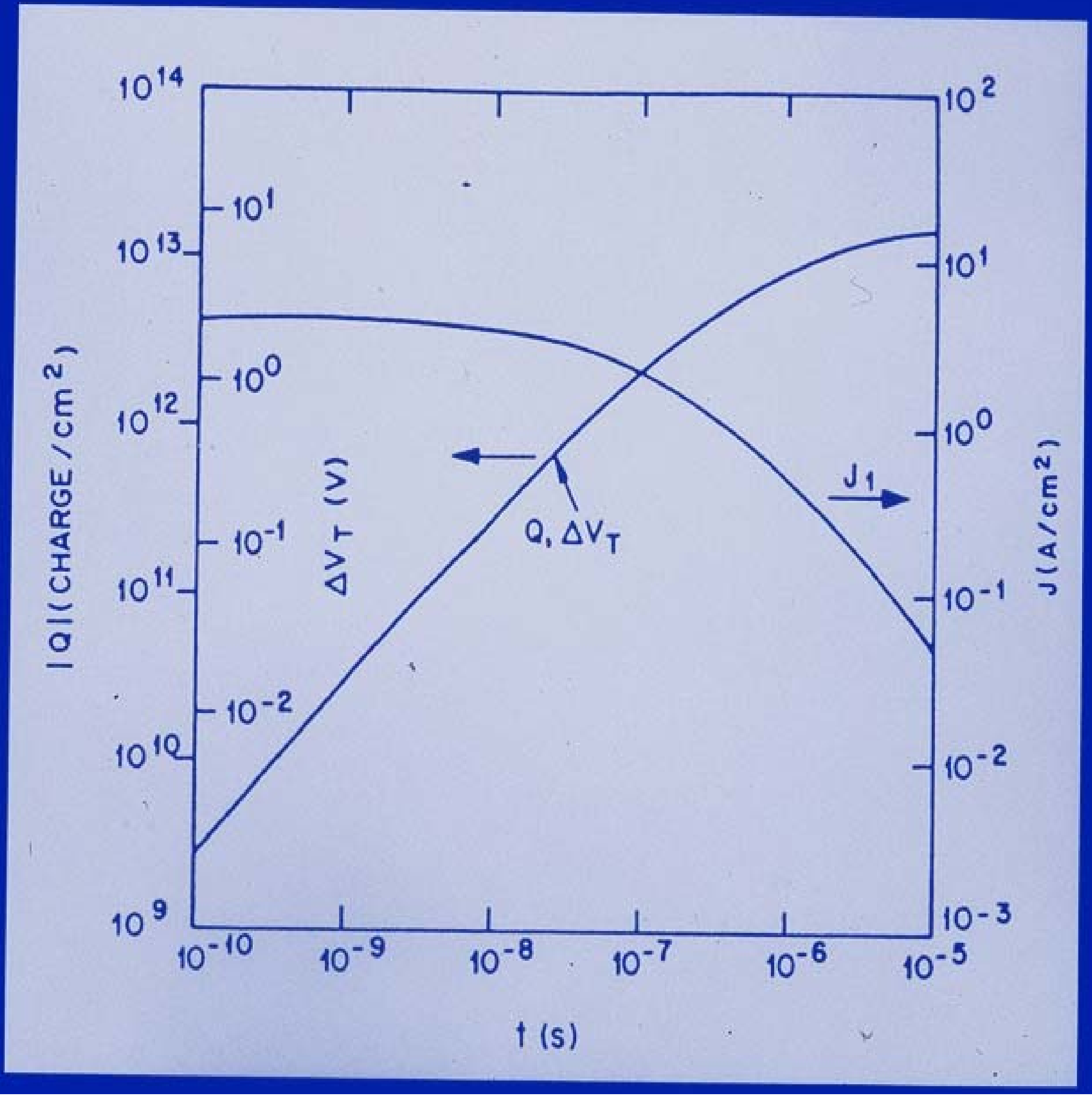


(a) Fowler-Nordhem Tunneling

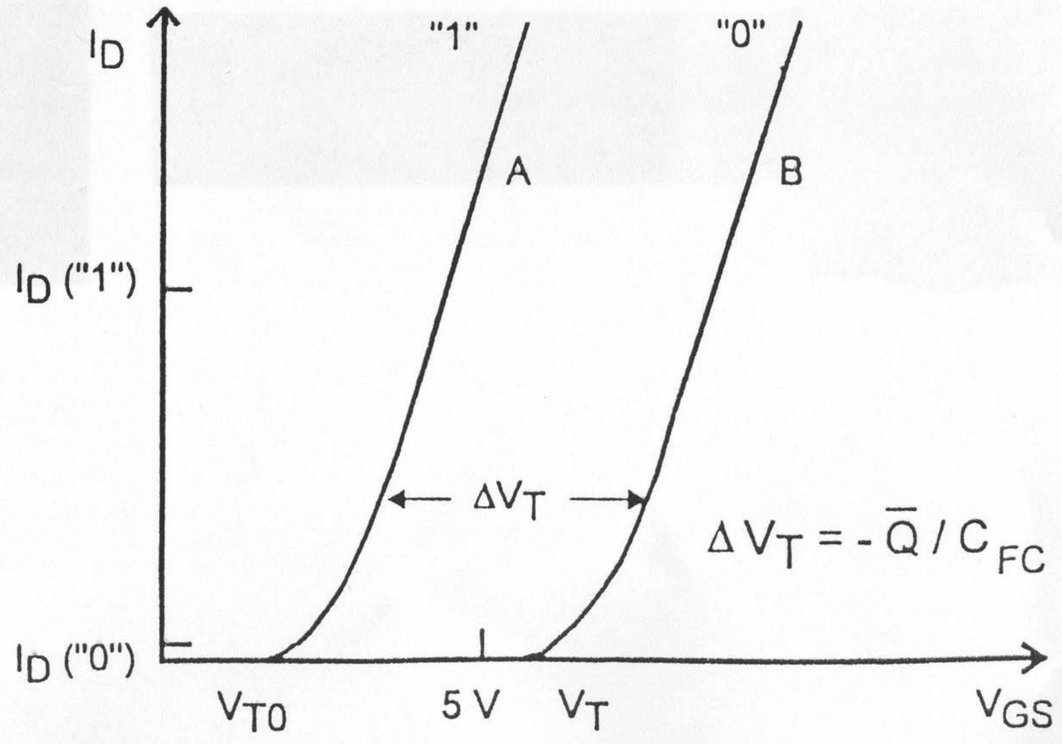


(b) Hot-Electron Injection

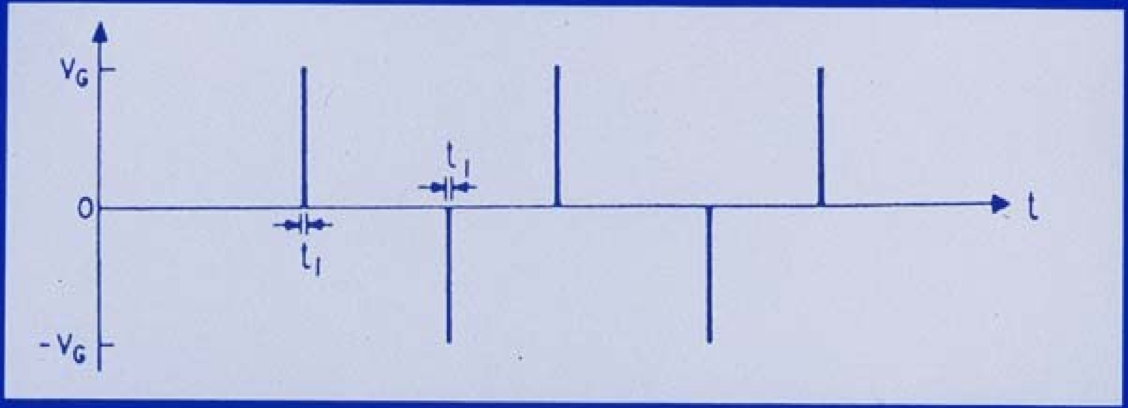
PROGRAMMING CURRENT AND STORED CHARGE



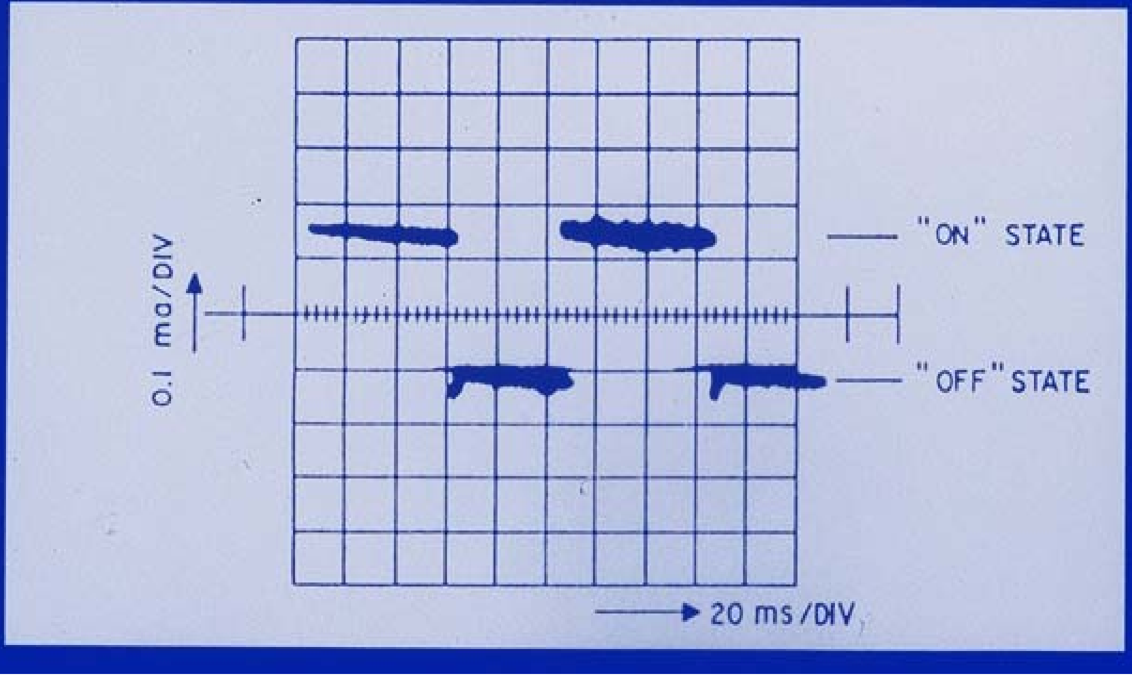
THRESHOLD VOLTAGE SHIFT DUE TO CHARGE STORAGE ON THE FLOATING GATE



DEMONSTRATION OF EEPROM OPERATION



(a) Applied gate pulse voltage

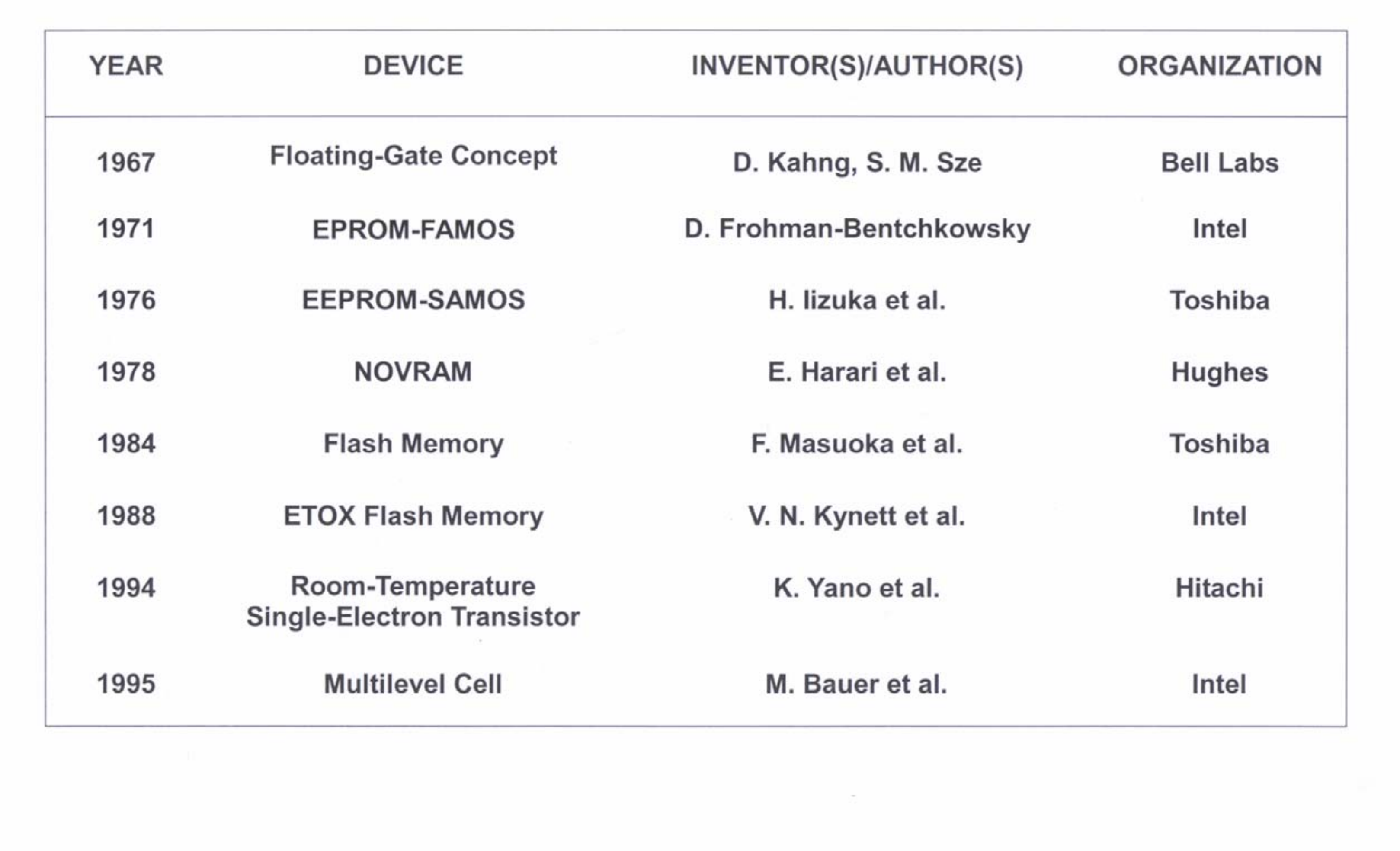


(b) Source-drain current
For $V_G = \pm 50$ V, $t_1 = 0.5$ µs

FIRST PAPER ON NVSM (1967)

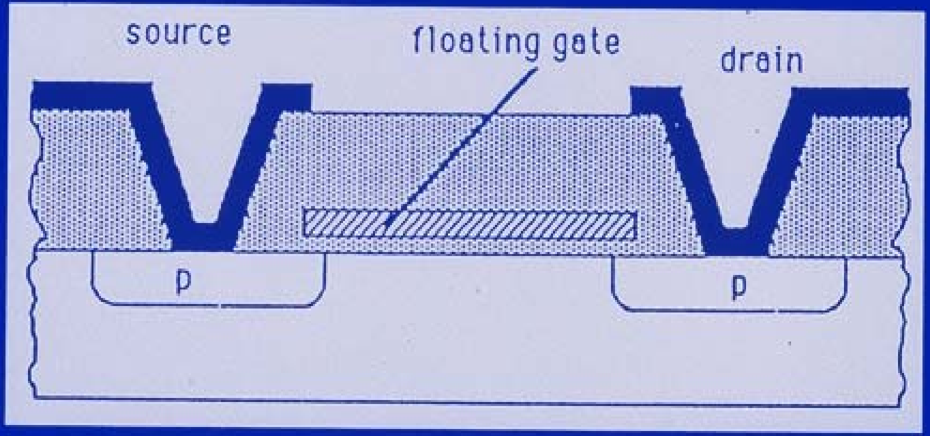
- The possibility of nonvolatile memory in semiconductors was recognized for the first time, and an experimental electrically erasable-programmable read only memory was demonstrated.
- The paper introduced not only the basic concept of nonvolatility, but also the floating-gate structure which has been the dominant technology for nonvolatile charge storage.
- The paper introduced nonlinear transport processes for programming and erase. These approaches have been adopted for most NVSM devices.

HISTORY OF FLOATING-GATE NVSM

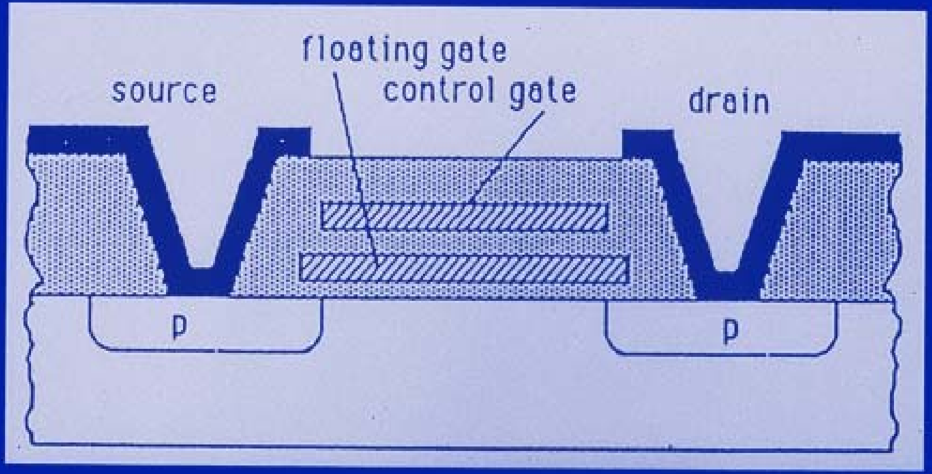


YEAR	DEVICE	INVENTOR(S)/AUTHOR(S)	ORGANIZATION
1967	Floating-Gate Concept	D. Kahng, S. M. Sze	Bell Labs
1971	EPROM-FAMOS	D. Frohman-Bentchkowsky	Intel
1976	EEPROM-SAMOS	H. Iizuka et al.	Toshiba
1978	NOVRAM	E. Harari et al.	Hughes
1984	Flash Memory	F. Masuoka et al.	Toshiba
1988	ETOX Flash Memory	V. N. Kynett et al.	Intel
1994	Room-Temperature Single-Electron Transistor	K. Yano et al.	Hitachi
1995	Multilevel Cell	M. Bauer et al.	Intel

TWO EARLY VERSIONS OF FLOATING GATE MEMORY

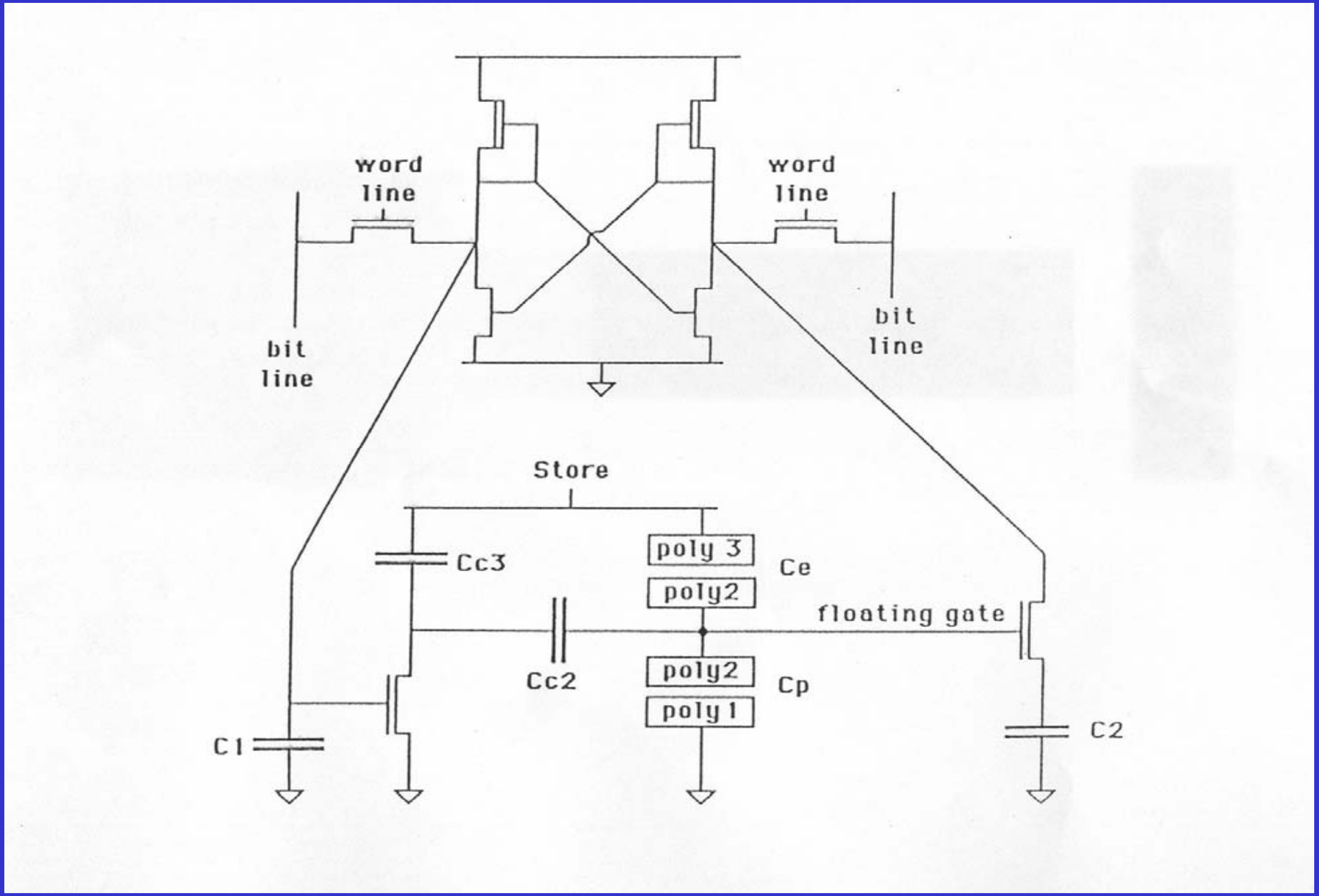


In 1971
FAMOS (Floating Gate
Avalanche Injection MOS)

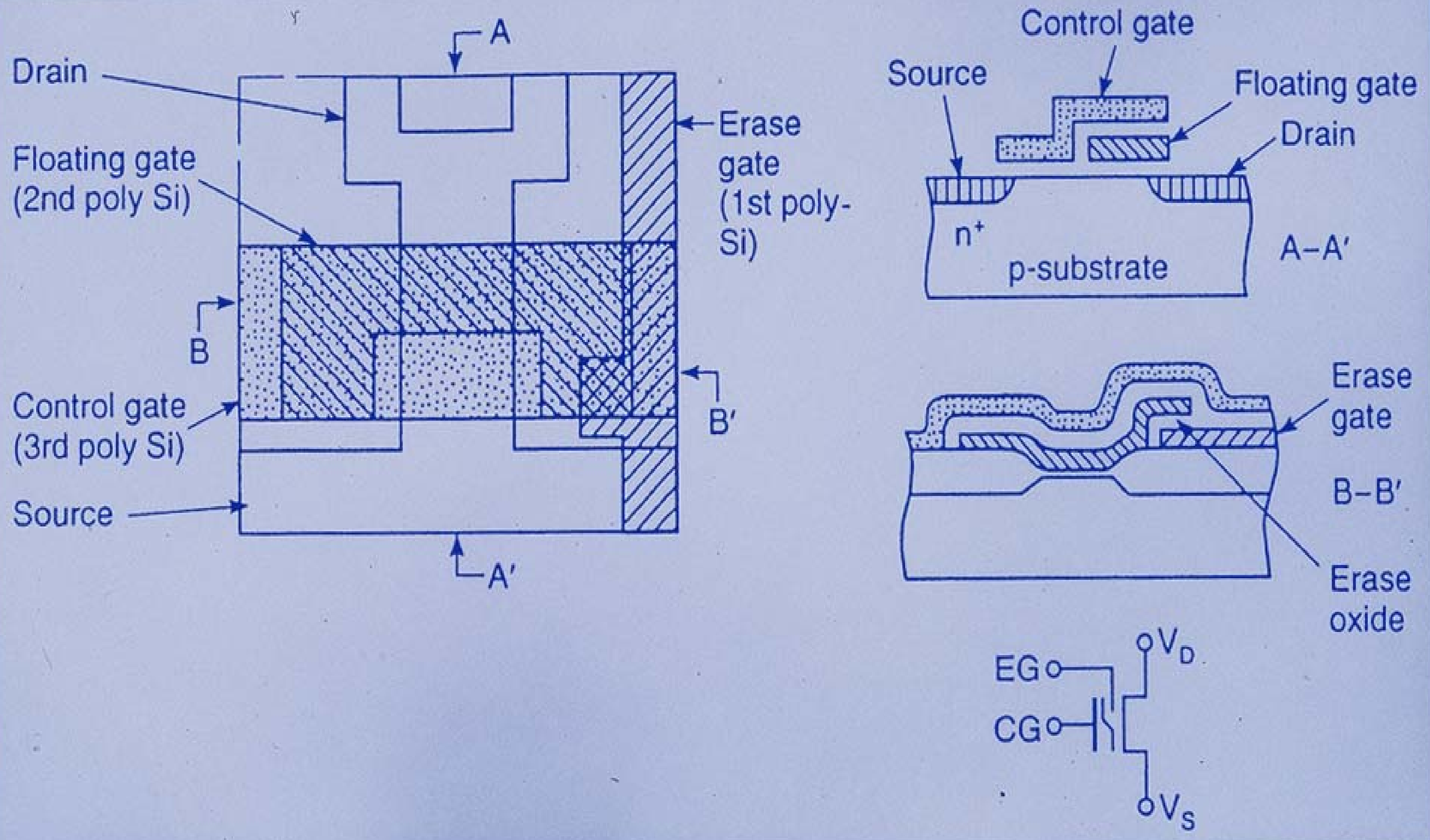


In 1976
SAMOS (Stacked Gate
Avalanche Injection MOS)

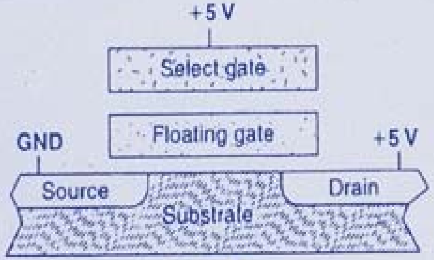
NOVRAM (Non-Volatile RAM) IN 1978



TOP AND CROSS-SECTIONAL VIEW OF A FLASH MEMORY (1984)

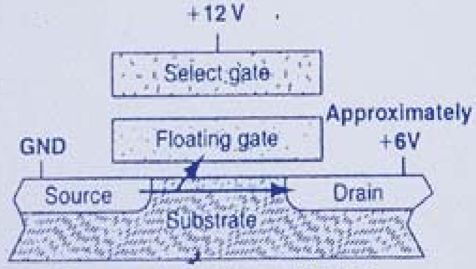


AN ETOX CELL OPERATION
(EPROM with Tunnel Oxide) in 1988



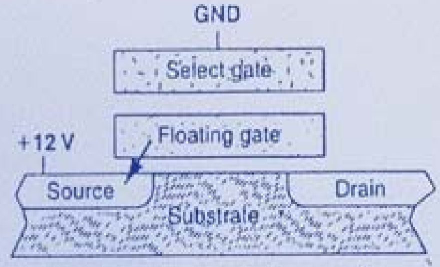
Read

(a)



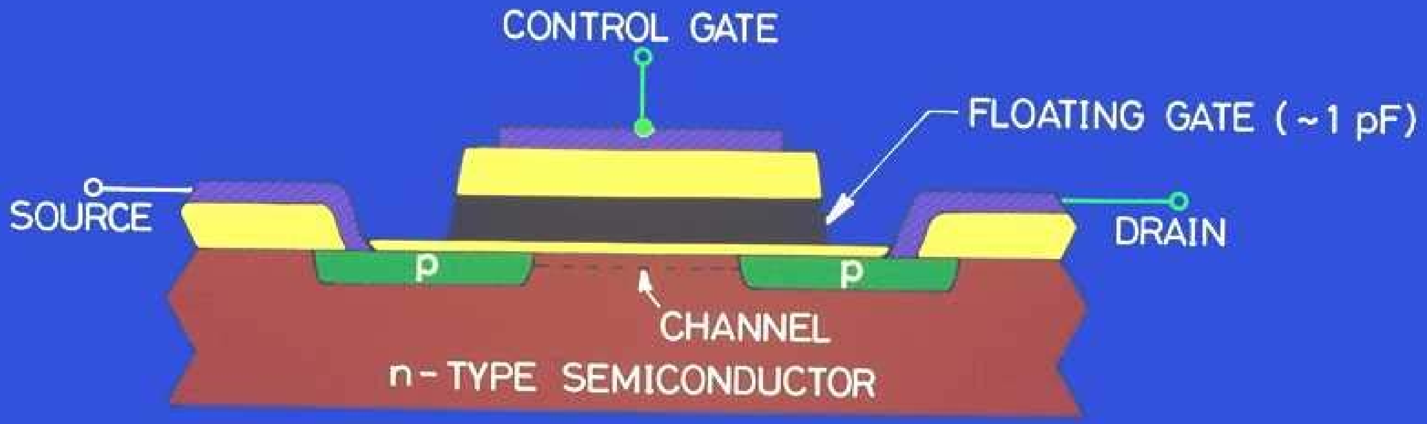
Program, hot electron injection

(b)

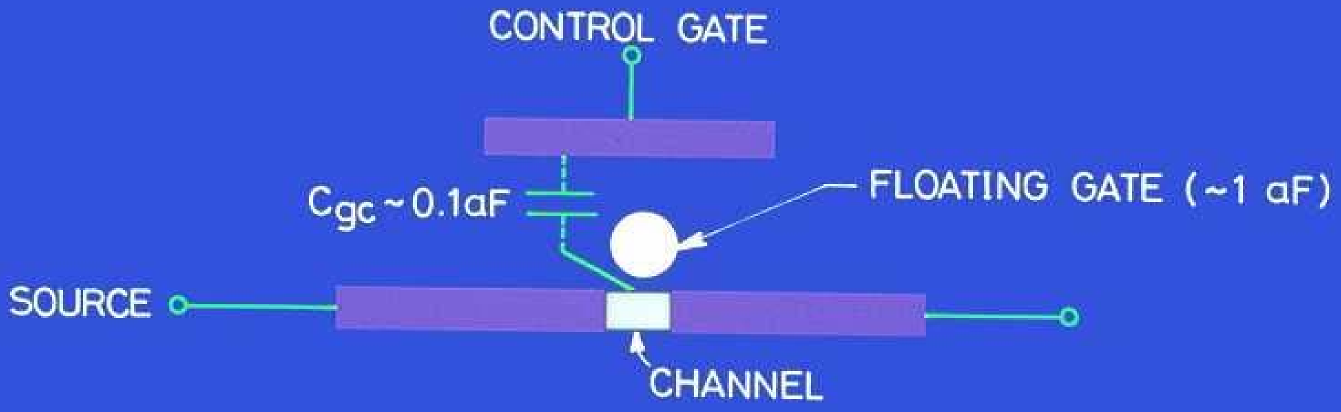


Erase, Fowler-Nordheim tunneling

SINGLE-ELECTRON MEMROY CELL
- A Limiting case of the Floating Gate NVSM

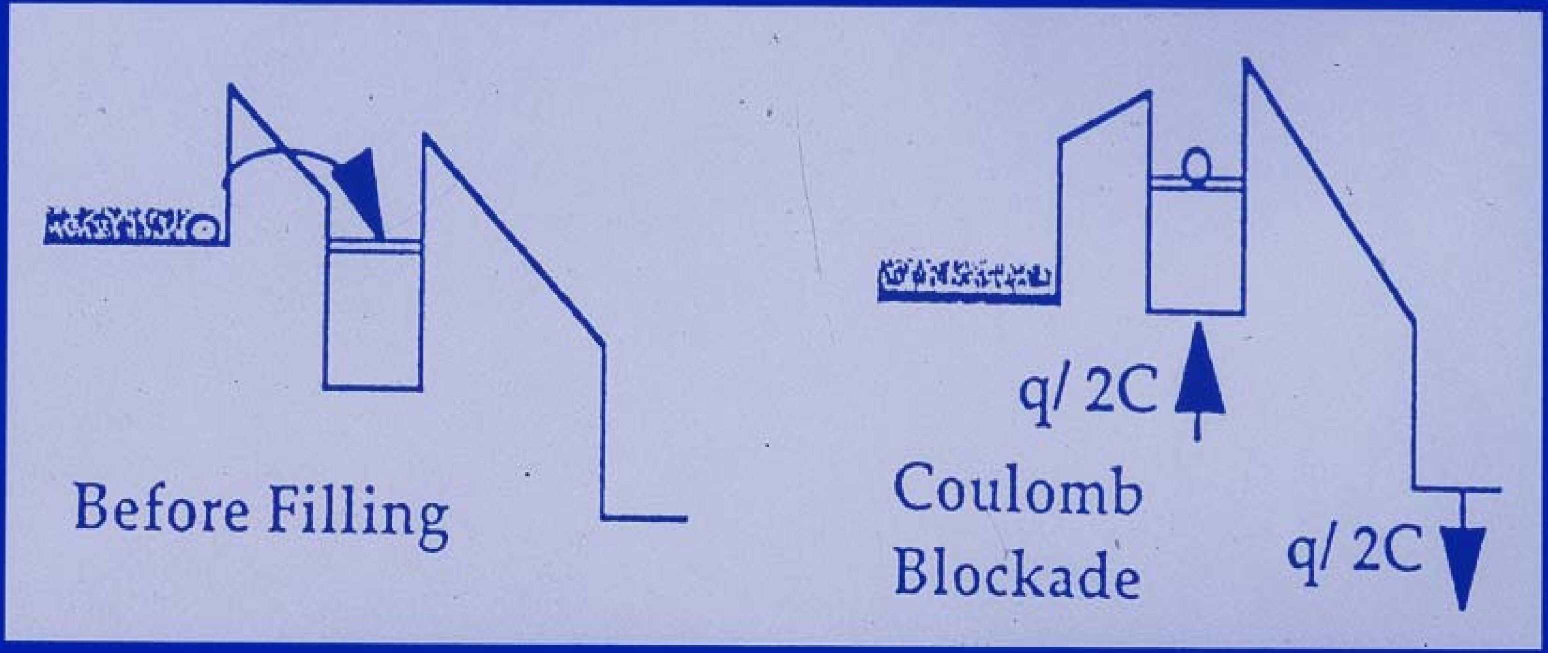


(a)



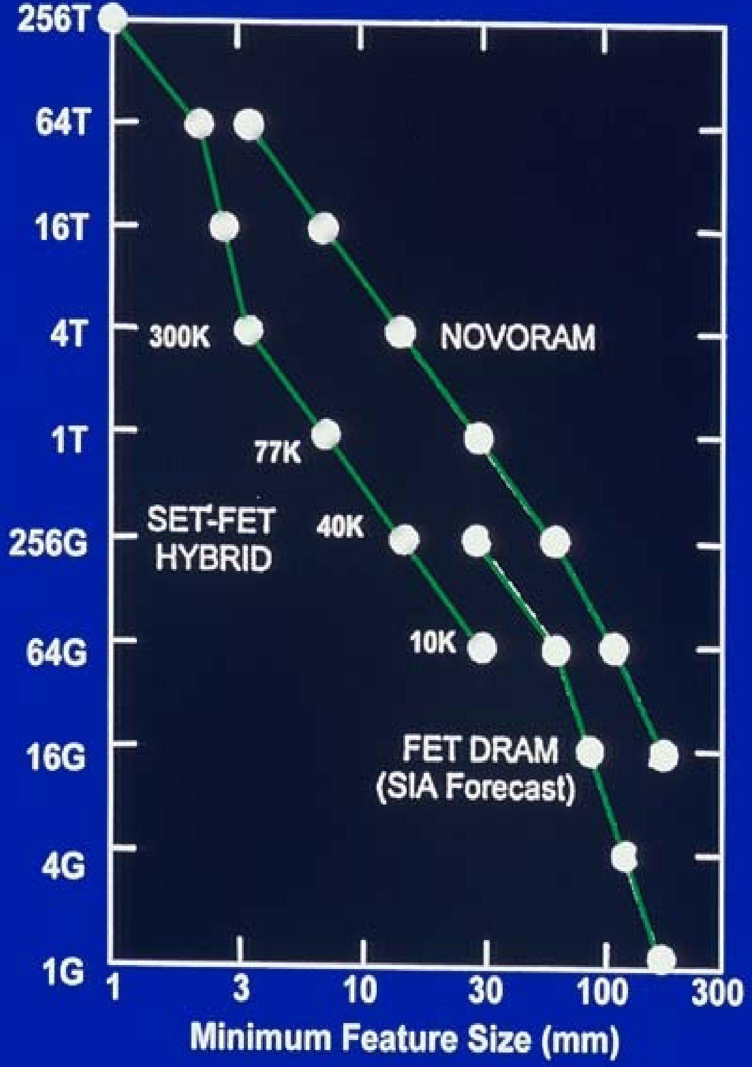
(b)

SINGLE – ELECTRON TRANSISTOR

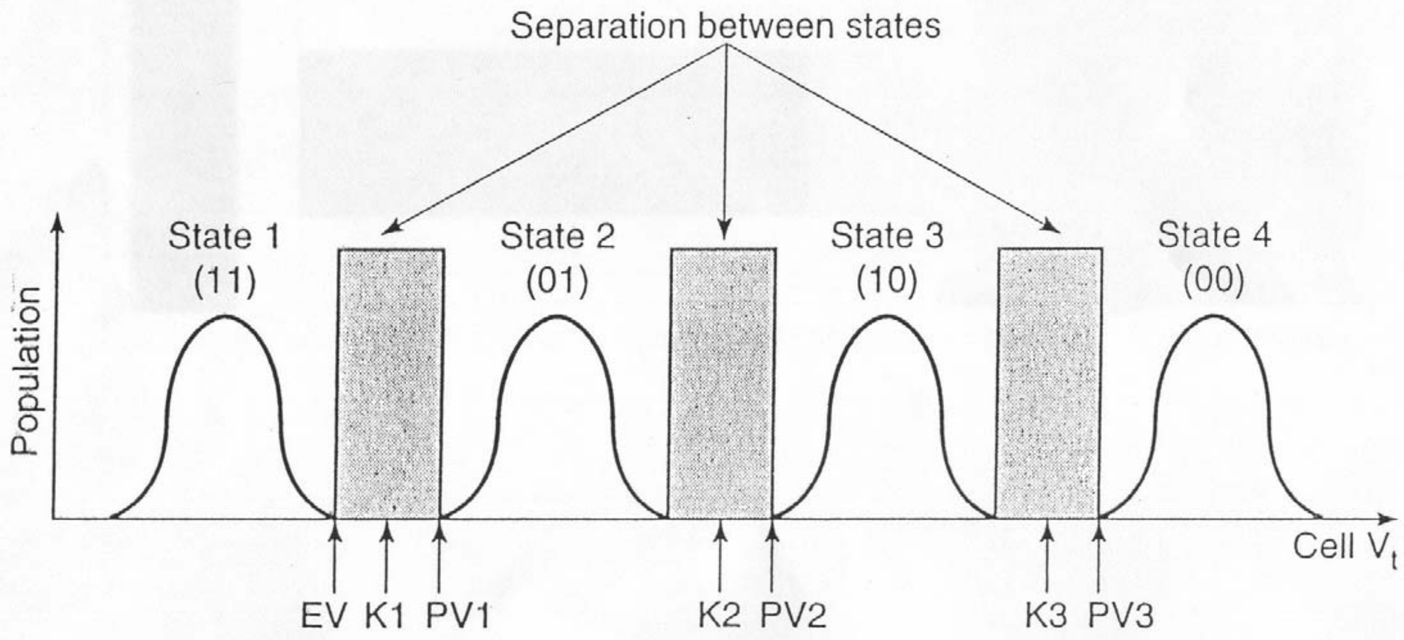


Coulomb Blockade: the effect of the electron reduces the potential in the quantum well and blocks the transfer of another electron

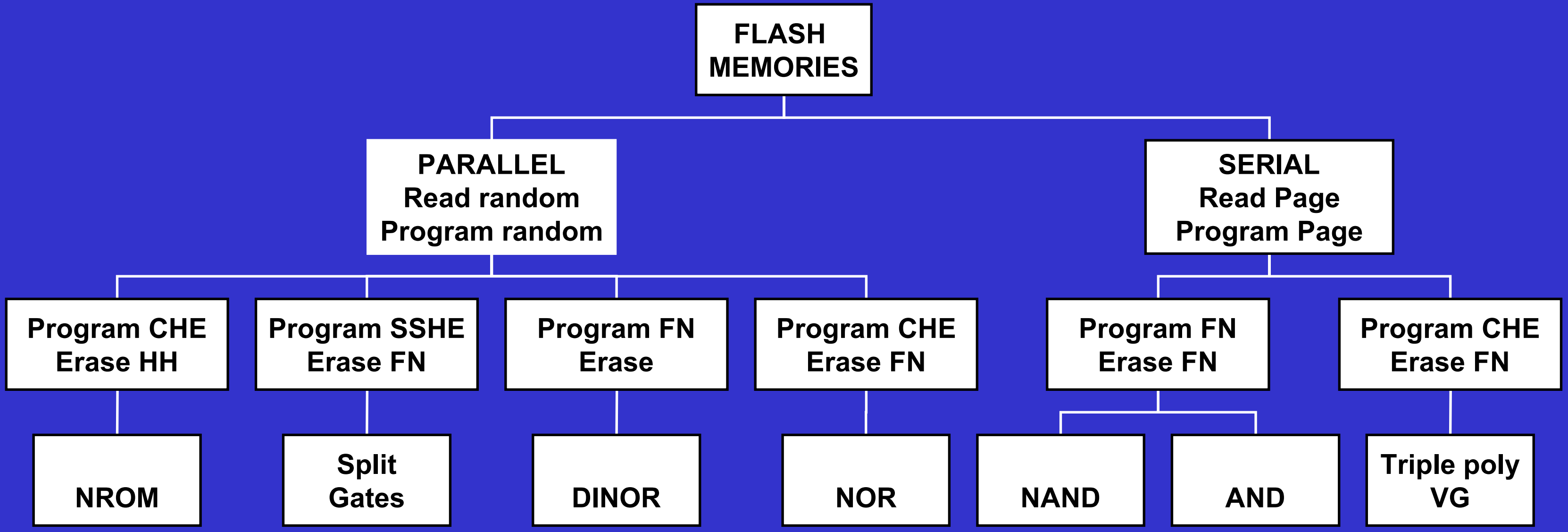
MEMORY DENSITY OF DRAM AND SET



MULTILEVEL-CELL
THRESHOLD VOLTAGE DISTRIBUTION



FAMILY TREE OF FLASH ARCHITECTURES



FLASH MEMORIES

PARALLEL
Read random
Program random

SERIAL
Read Page
Program Page

Program CHE
Erase HH

Program SSHE
Erase FN

Program FN
Erase

Program CHE
Erase FN

Program FN
Erase FN

Program CHE
Erase FN

NROM

Split
Gates

DINOR

NOR

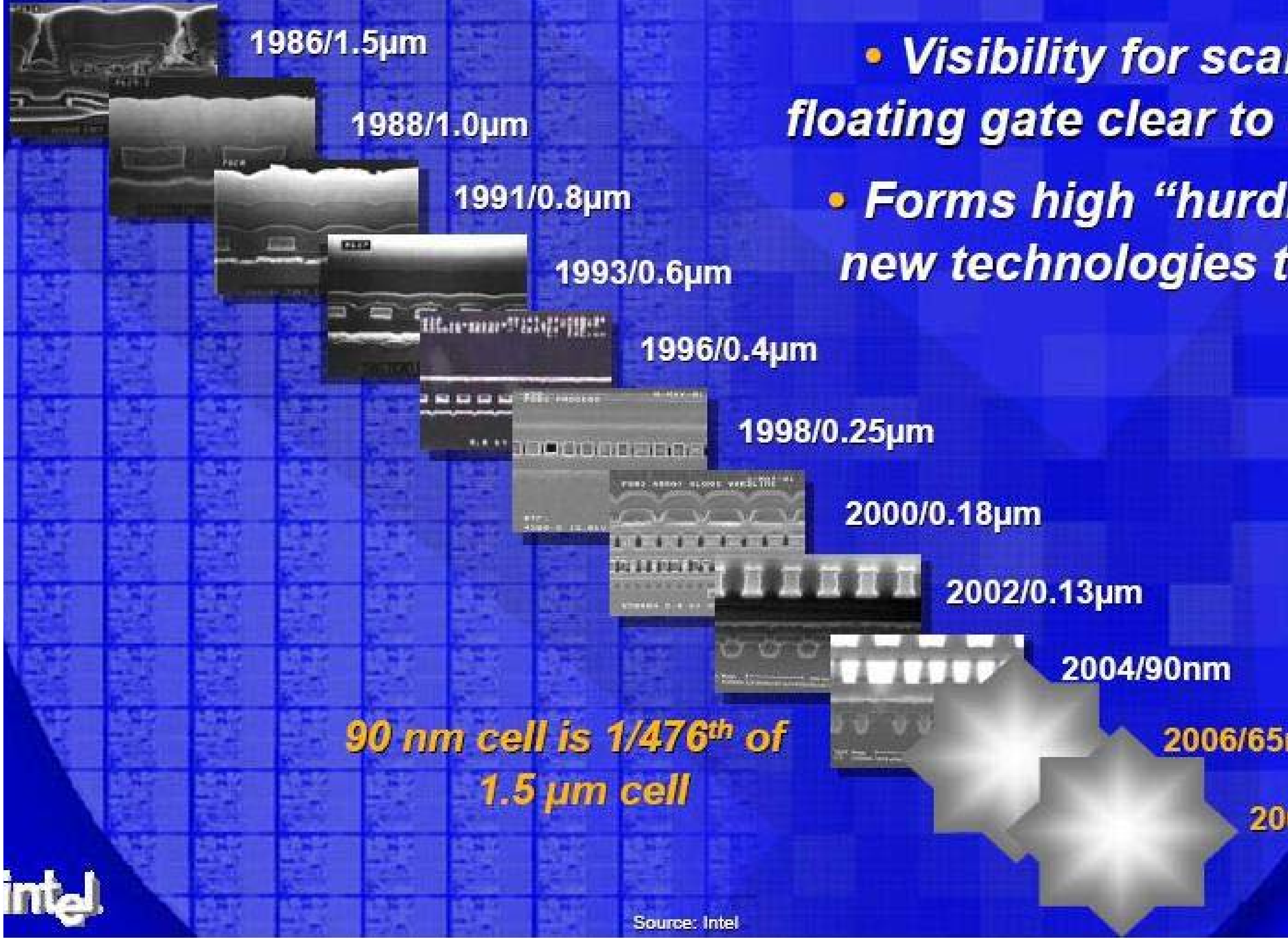
NAND

AND

Triple poly
VG

"Moore's Law" Continues with FG NVM

- *Visibility for scaling of floating gate clear to 32 nm*
- *Forms high "hurdle" for new technologies to leap*

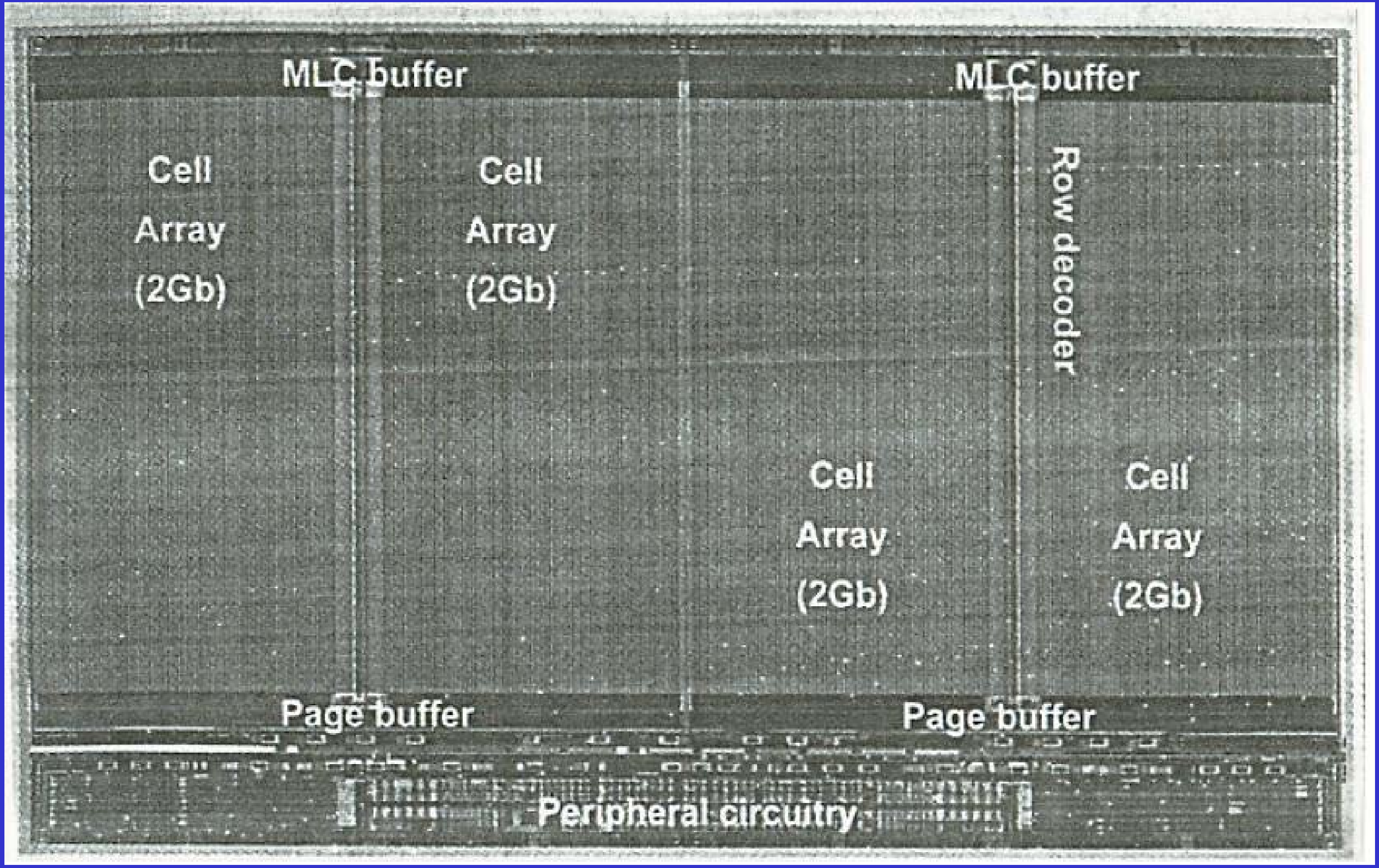


90 nm cell is 1/476th of 1.5 µm cell

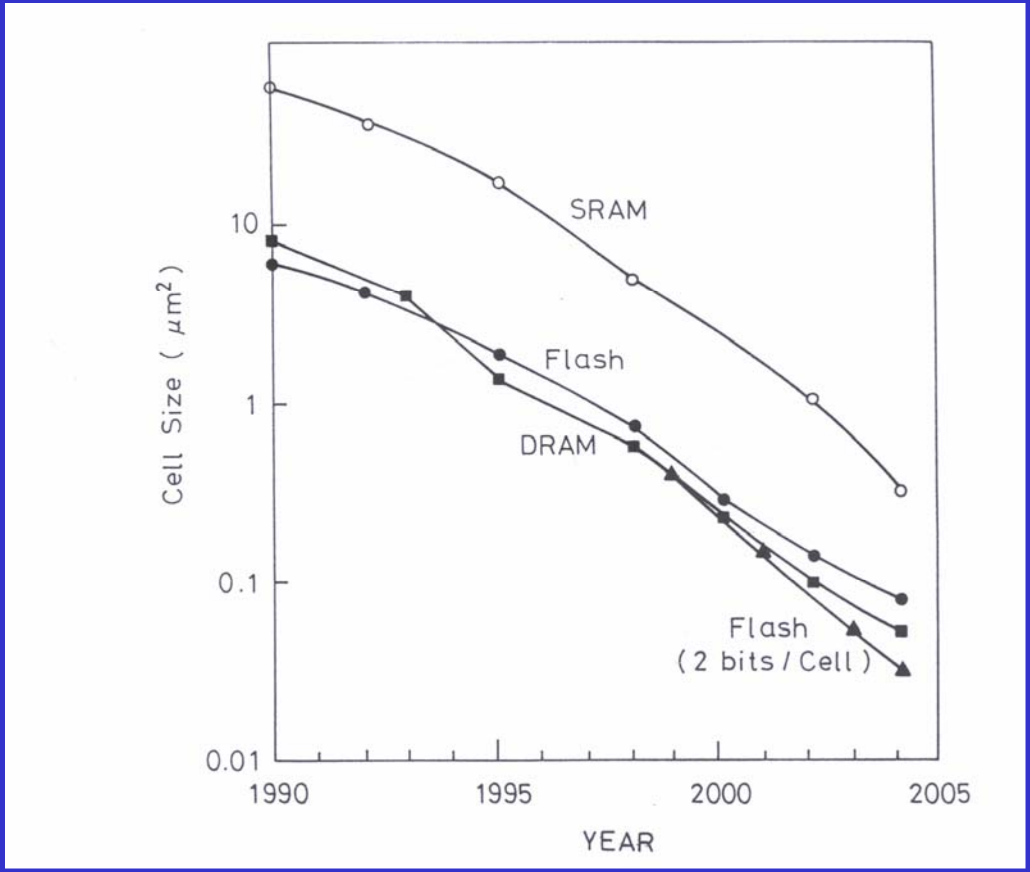
Source: Intel

8 Gb FLASH MEMORY

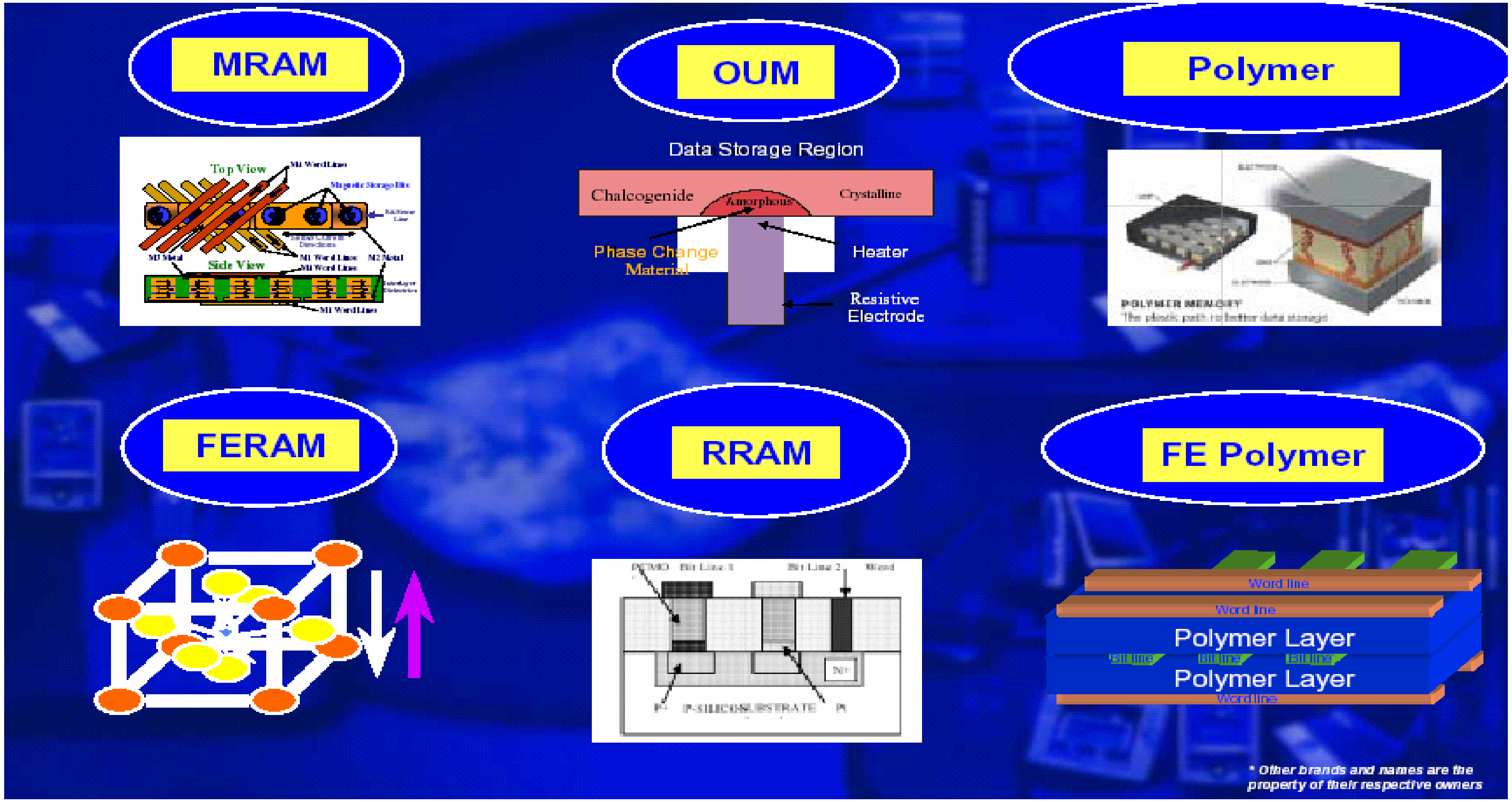
(3 V 65 nm 133 mm^2 0.02μm^2/cell)



CELL SIZE REDUCTION VERSUS YEAR



EMERGING NONVOLATILE MEMORIES



* Other brands and names are the property of their respective owners

EMERGING NONVOLATILE MEMORIES

- MRAM (Magneto-resistive Random Access Memory): based on tunneling Magneto-resistance effect.
- FeRAM (Ferroelectric RAM) : based on remanent polarization in Pervoskite materials.
- OUM (Ovonies Unified Memory) also called PCRAM (Phase-change RAM): based on reversible phase conversion between the amorphous and the crystalline state of a chalcogenide glass, which is accomplished by heating and cooling of the glass.
- RRAM (Resistance RAM): based on change in resistance with applied electric field.
- Polymer Memory: based on resistance change of polymer at cross point of metal layers.
- Millipede Memory: based on concept similar to punch cards, using thermally assisted, rewritable displacement media such as PMMA.
- Nanocrystal Memory: based on distributed nano-floating-gate concept.

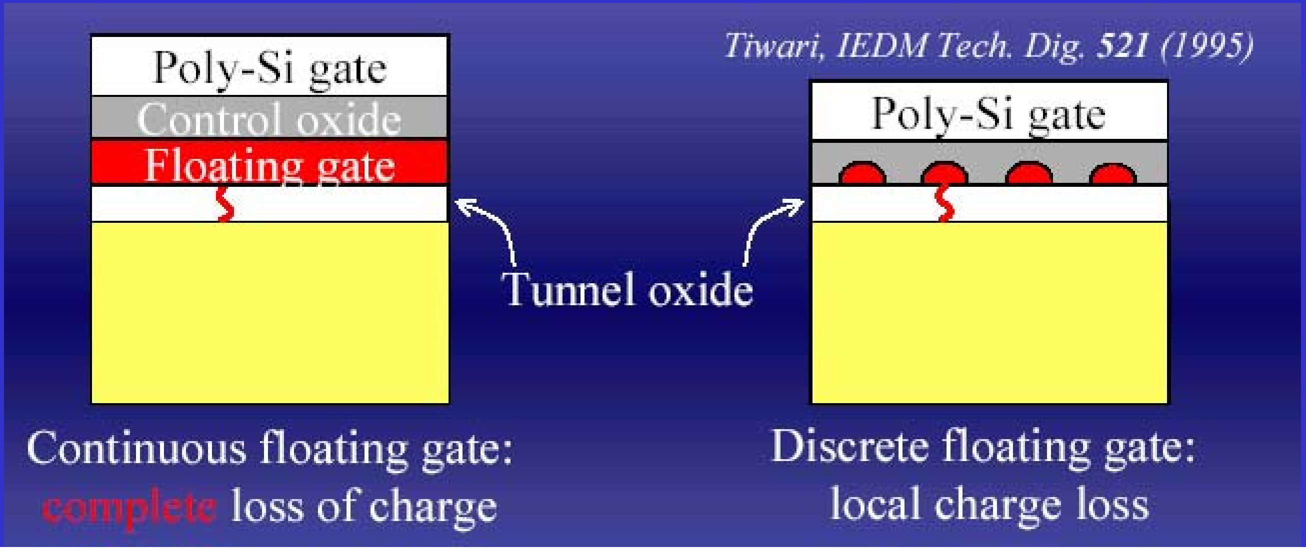
NANOCRYSTAL MEMORY

- Nanocrystal memory is a multiple-nano-floating-gate structure. Instead of charge stored in a single floating gate, charge is stored in many floating nanocrystals.

- Nanocrystal memory can extend the floating-gate scaling by allowing a further decrease in the tunnel oxide thickness.

- Nanocrystal memory offers enhanced robustness and fault-tolerance of distributed charge storage.

NANOCRYSTAL MEMORY



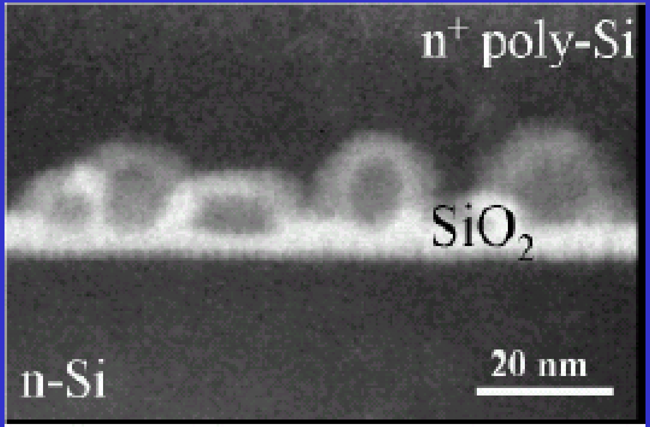
Tiwari, IEDM Tech. Dig. ***521*** *(1995)*

Continuous floating gate: complete loss of charge

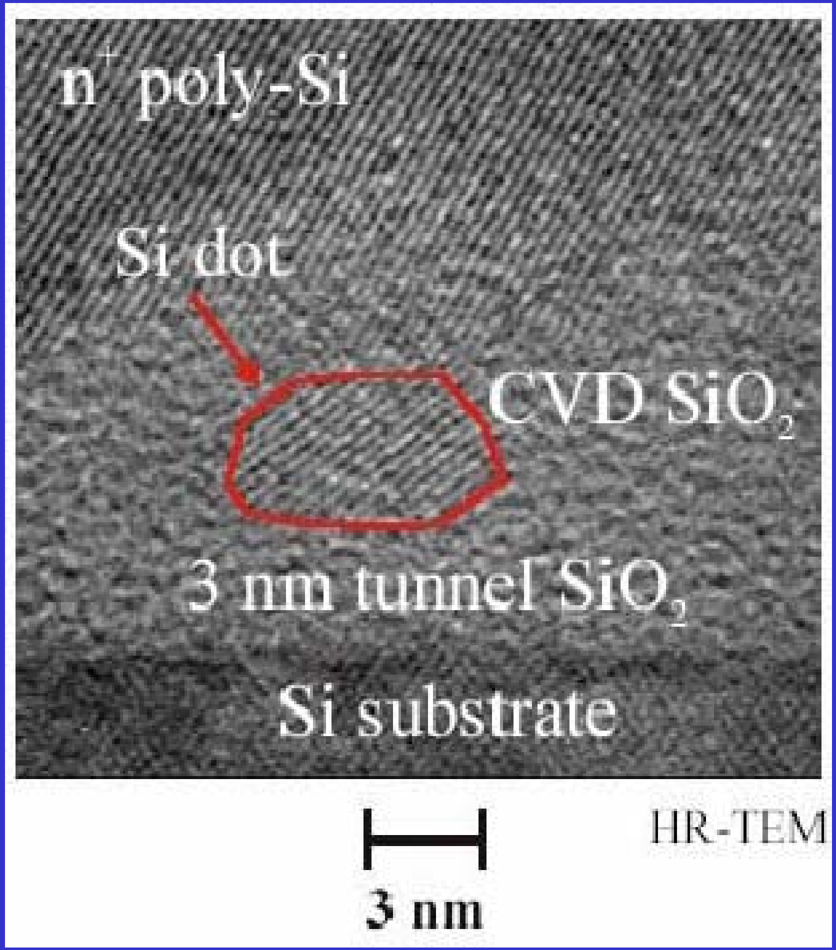
Discrete floating gate: local charge loss

Advantage of discrete storage nodes: robustness to stress induced leakage current ⇨ scaling of the tunnel oxide

TEM OF NANOCRYSTAL MEMORY



EFTEM tuned to the
SiO_2 plasmon energy



NANOCRYSTAL FABRICATION BY POLYMER SELF ASSEMBLY

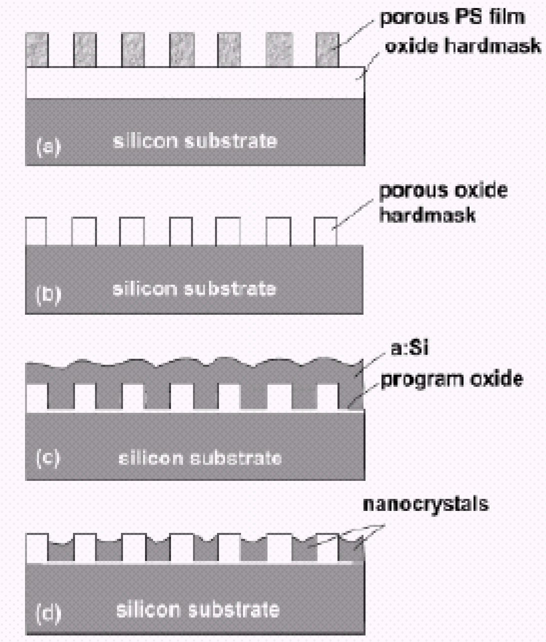


Fig. 1. (a-d) Schematic diagrams of the process flow used to create an array of Si nanocrystals beginning with a porous self assembled polymer template.

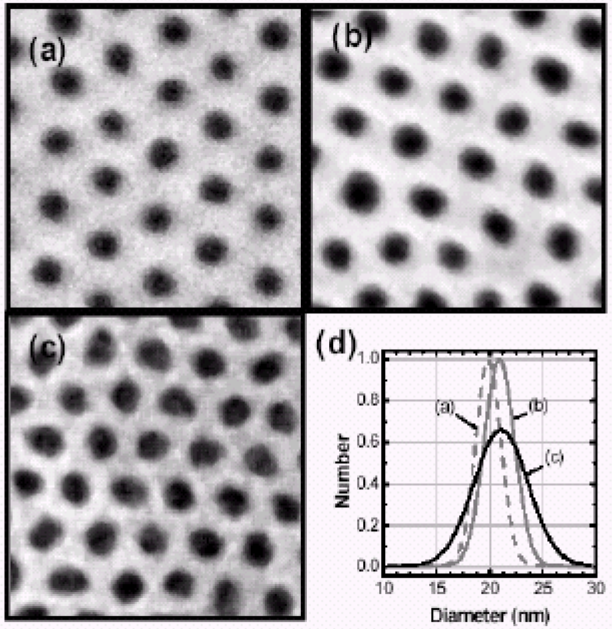


Fig. 2. (a-c) 200×200 nm top-down SEM images of (a) porous polymer film on SiO_2. (b) porous dielectric after pattern transfer into the SiO_2, and (c) Si nanocrystal array. (d) Distribution of polymer pore, dielectric pore, and Si nanocrystal diameters.

"Low voltage, scalable nanocrystal FLASH memory fabricated by templated self assembly", K.W. Guarini et al., 2003 IEDM Technical Digest, Session 22.2.

FLASH APPLICATIONS

MEMORIES FOR MOBILITY

- Nonvolatile = low power: power down to save power
- Low cost: one memory for program and data storage
- ≠magnetic hard disk: need solid state, low power, rugged
- Small form factor and light weight

MOBILE LIFESTYLE : DATA ANYTIME, ANYWHERE

APPLICATIONS OF NVSM

- **Communications**

  Cellular Phone, Cordless, Bluetooth, Pagers, Modems, Internet Appliance, Line Card, PBX, Set Top Box, LAN Modules, Network System, Network Adapters, Data Communication (5.6 Billion Units from 1999 to 2004).

- **Computer and Peripheral**

  PC, Flash-Drive Notebook, USB Drive, Personal Digital Assistant, Hard-Disk Drive, Portable DVD, Graphics Card, FAX, Printer (3.5 Billion Units).

- **Consumer**

  Digital Camera, Digital Voice Recorder, MP3 Music Player, Digital Camcorder, Telephone Answering Device, Games, Dictionary, Organizer, Toy, Electronic Book, Directories (2.5 Billion Units).

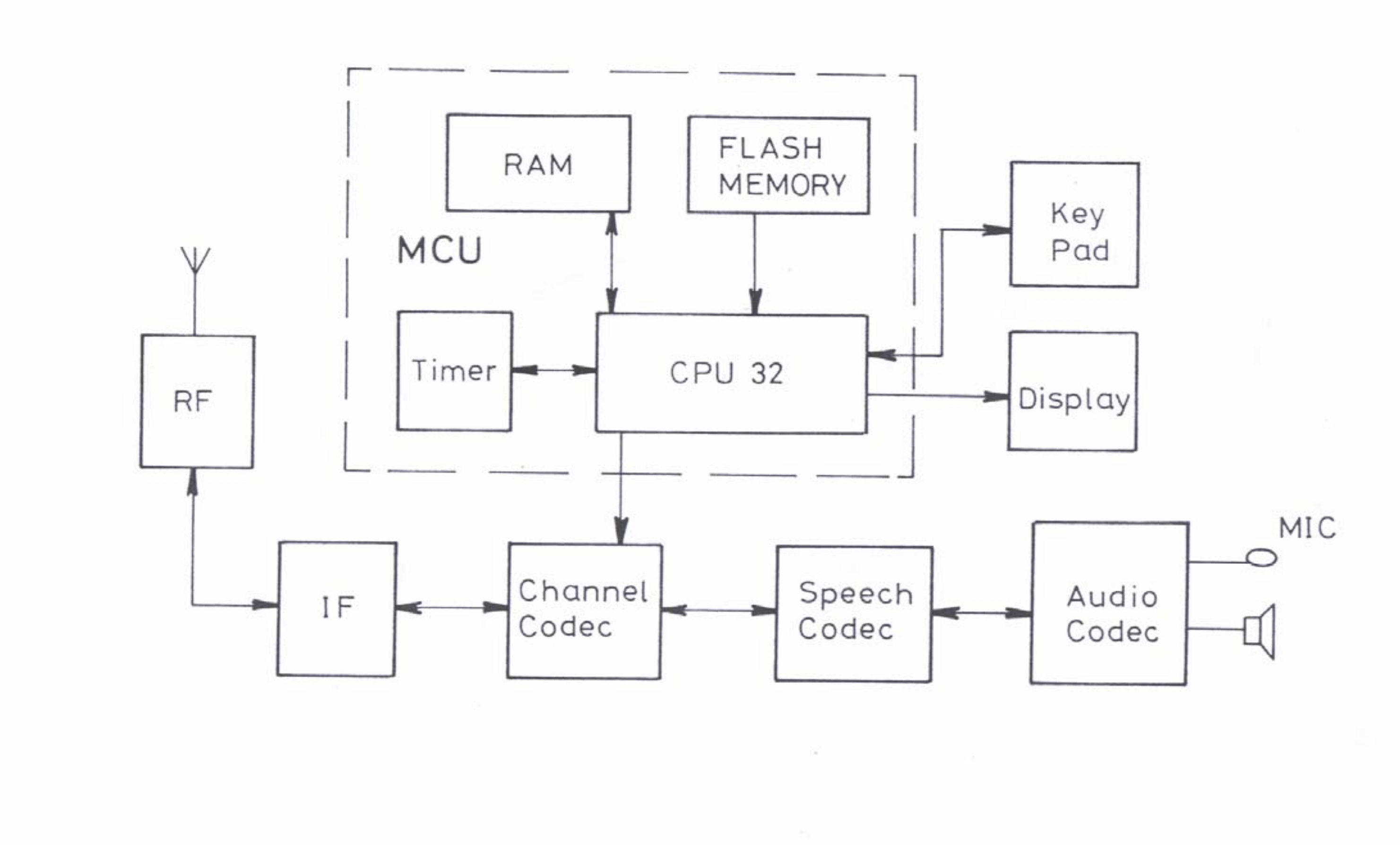
- **Transportation**

  Automotive systems, Global Positioning System (Auto, Marine, Aviation) (570 Million Units).

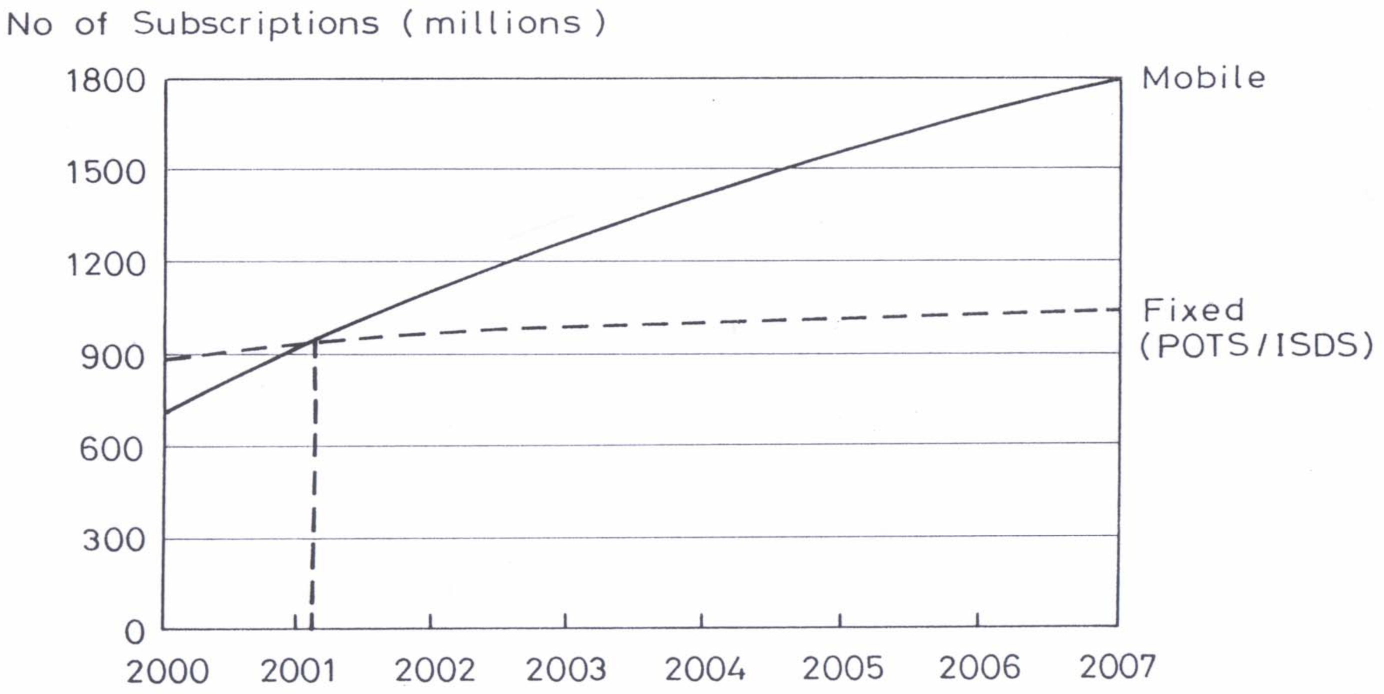
- **Industrial**

  Meter, Sensing Device, Embedded System, POS, Solid State Drive (630 Million Units).

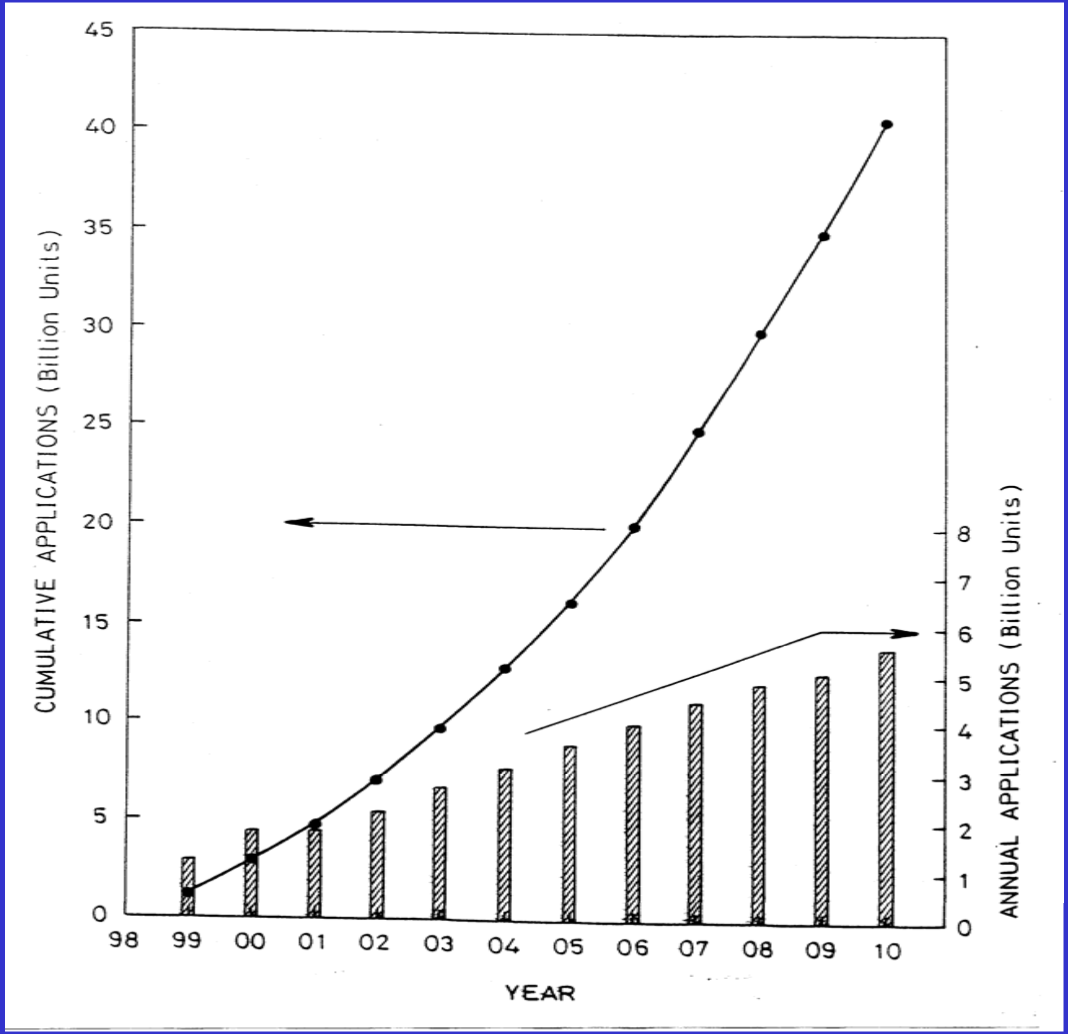
CELLULAR PHONE BLOCK DIAGRAM



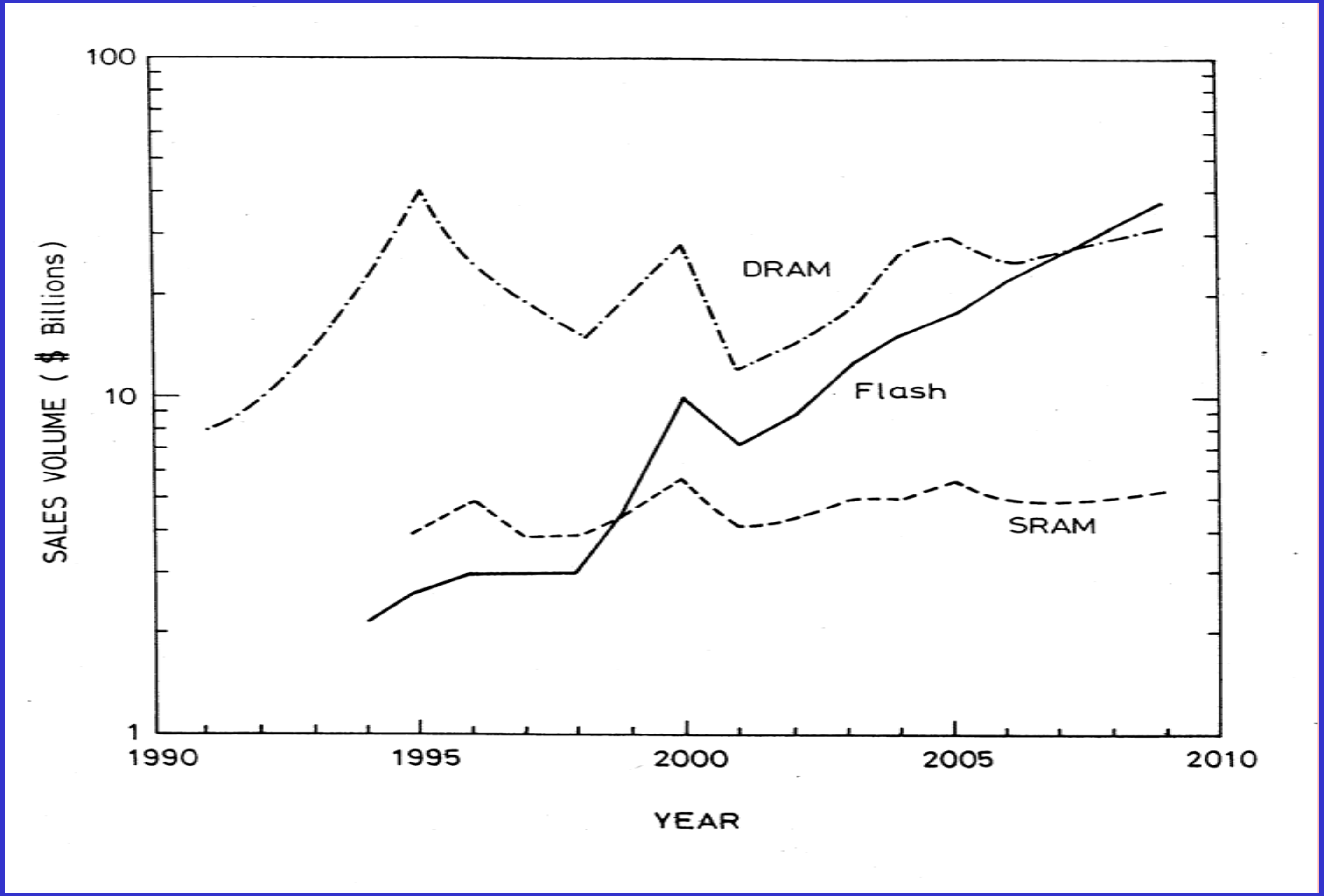
GLOBAL TELEPHONE SUBSCRIPTIONS



END-USE APPLICATIONS OF NVSM (1999-2010)



SEMICONDUCTOR MEMORY MARKET



TECHNOLOGY DRIVER FOR ELECTRONIC INDUSTRY

(Application Created by the New Technology)

CONCLUSION

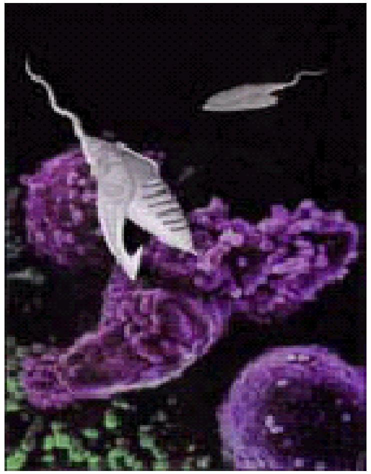
- The nonvilatile semiconductor memory (NVSM) was invented in 1967 based on the floating-gate concept for charge storage and nonlinear transport processes for programming and erase.
- Because of its attributes of high density, low-power consumption, nonvolatility, and electrical rewritability, NVSM has surpassed SRAM in global memory market in 1999, and is poised to eclipse DRAM in the near future.
- NVSM has revolutionized electronic technology and enabled tremendous advances in portable electronic systems such as the mobile phone, notebook computer, digital photography, PDA, GPS, Smart IC card, etc.
- Many emerging NVSM technologies have been investigated to extend the floating-gate scaling to sub-10 nm, regime. A major candidate is the nanocrystal memory.

ACKNOWLEDGEMENTS

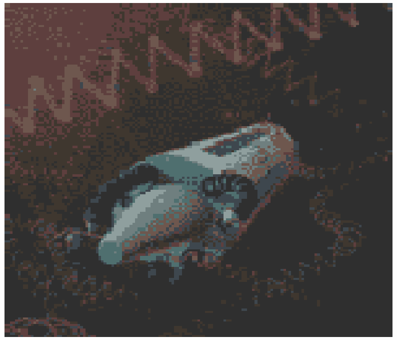
R. Bez et al	ST Microelectronics
Stefan Lai	Intel
S. Lombardo et al	CRN-IMM
F. Masauoka et al	Toshiba
Sang Lyul Min	Seoul University
A. Niebel	Web-feet Research

UNCONVENTIONAL APPLICATIONS OF NVSM

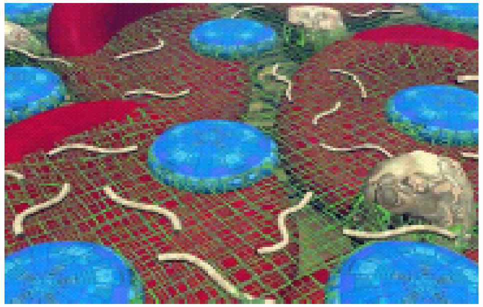
- Medical nanorobots



Destrobot



DNA Repair Machine

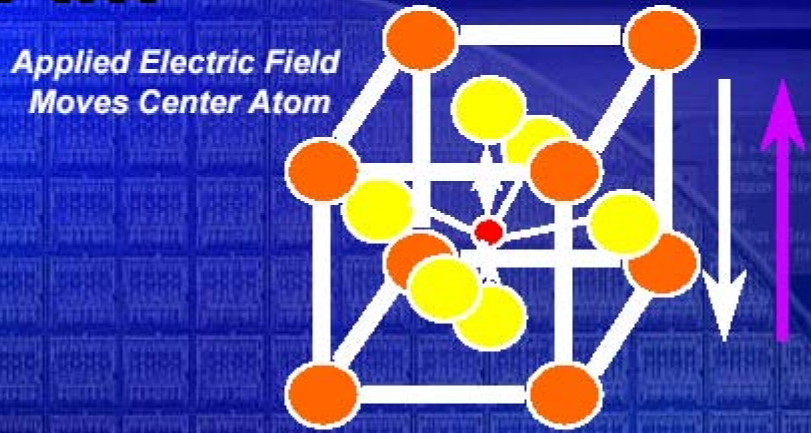


Digital DNA

- Sensor networks in general and battlefield applications in particular

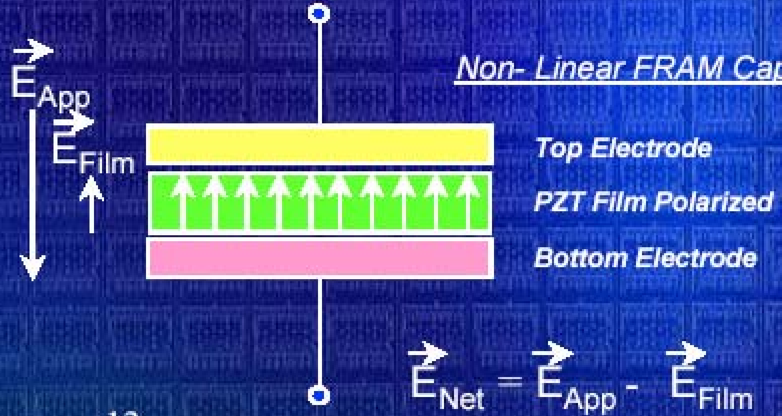
FeRAM

- Operation
  - Selected PZT crystalline materials have bi-stable center atom
  - Data is stored by applying an voltage to polarize the internal dipoles “Up” or “Down”
  - Reading by sensing the displacement current
  - Fast read/write < 100 nSec
  - Destructive read, limited number of read cycles due to fatigue
  - Very low power consumption



Perovskite Crystal Unit Cell
PZT (PbO,ZrO_2, TiO_2) Lead-Zirconate-Titanate

● Tetra/Pentavalent Atom
● Di/Monovalent Metal Atoms
● Oxygen Atoms



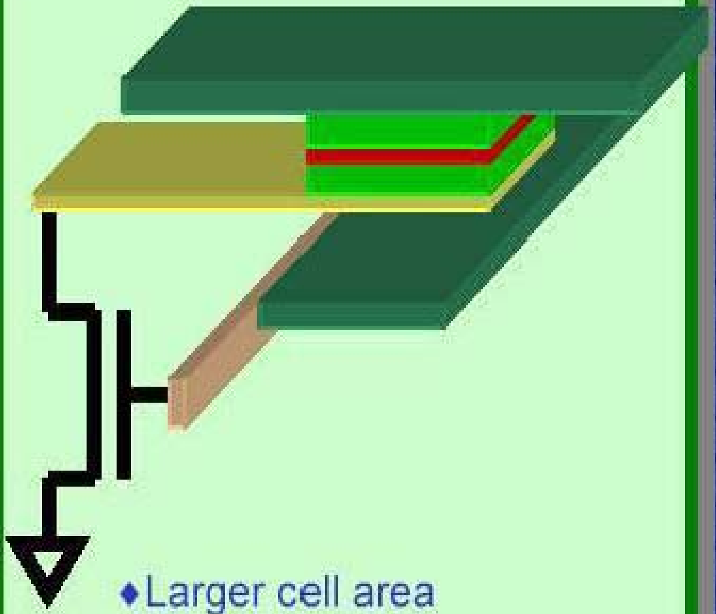
Non- Linear FRAM Capacitor

$$\vec{E}_{Net} = \vec{E}_{App} - \vec{E}_{Film}$$

MRAM

- Operation
  - Cell is 1 MJT + 1 Transistor
  - Electric current switches the magnetic polarity of sense layer
  - Change in magnetic polarity sensed as resistance change in tunnel junction
  - Non destructive read
  - Very fast read/write performance, < 10 nSec reported
  - Unlimited number of R/W endurance
  - Relatively high write current

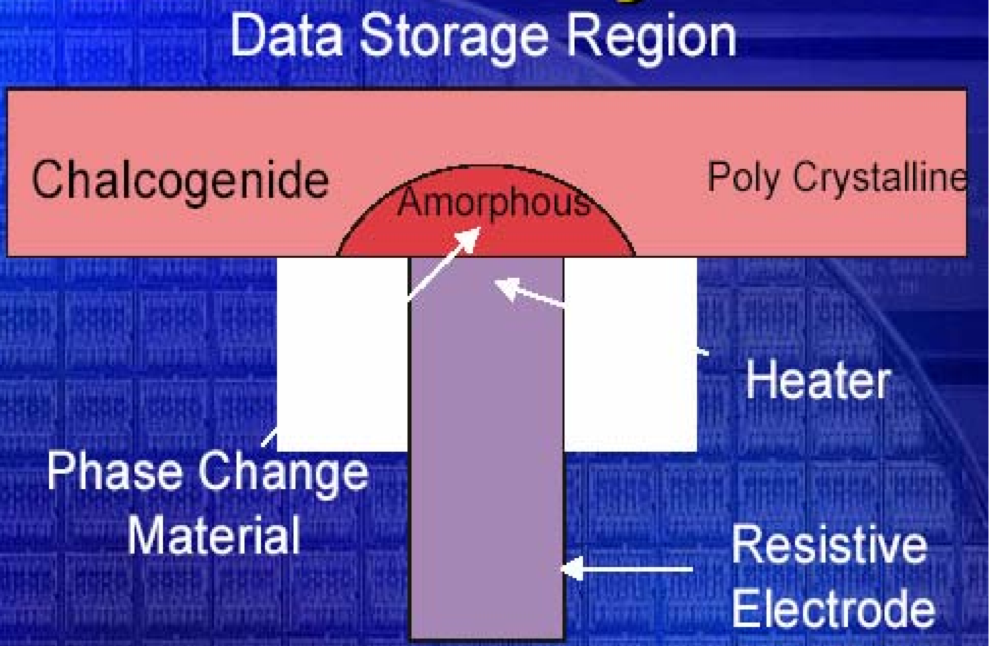
FET cell (1TJ-1T)

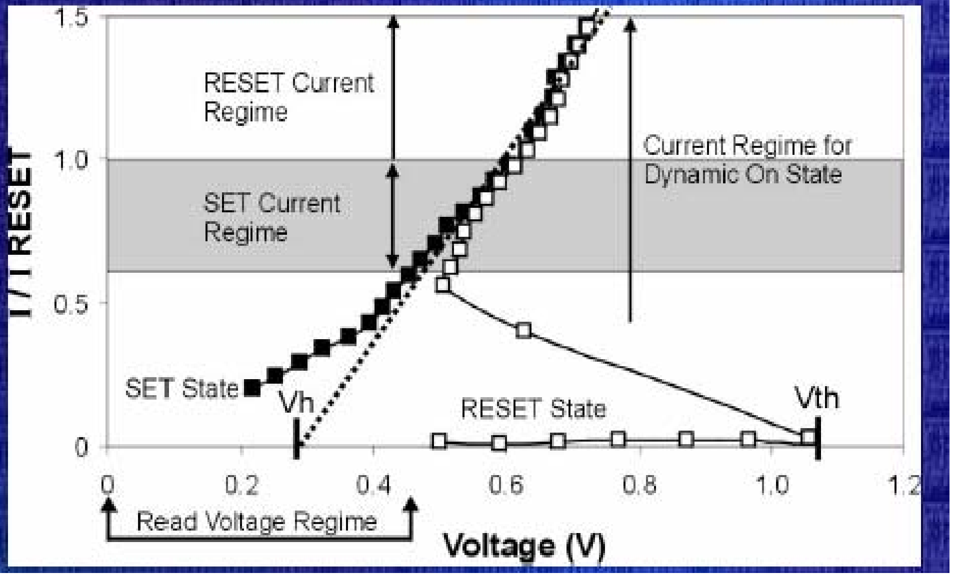


- Larger cell area (>8 F^2)
- Faster random access time
- SRAM like ($\leq$10ns)

Ovonics Unified Memory

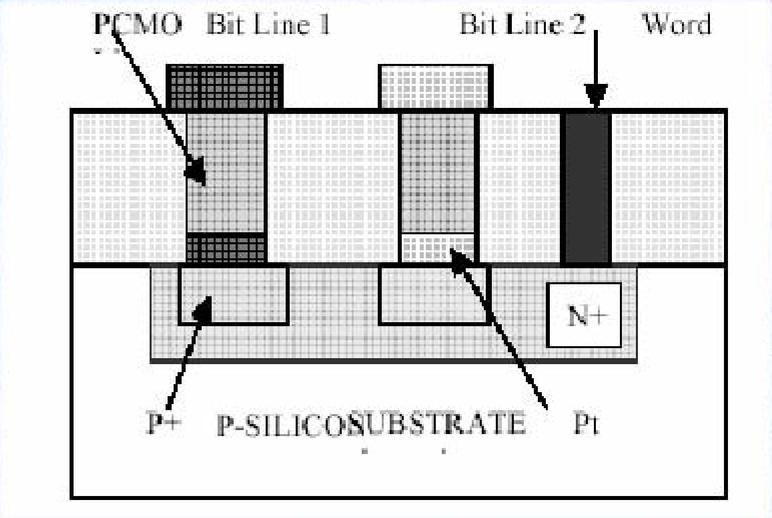
- Operation
  - Chalcogenide material alloys used in re-writable CDs and DVDs
  - Electrical energy (heat) converts the material between crystalline (conductive) and amorphous (resistive) phases
  - Cell reads by measuring resistance
  - Non-destructive read
  - ~10^{12} write/erase cycles
  - Medium write power
  - Relatively easy integration with CMOS





Complex Metal Oxide RRAM

- Operation
  - PCMO material, Complex metal oxide studied for high temp superconductivity
  - Change in resistance with applied electric field
  - Low resistance with forward bias, high resistance when reverse electric field applied
  - Low field read by measuring resistance
  - Relatively low power write



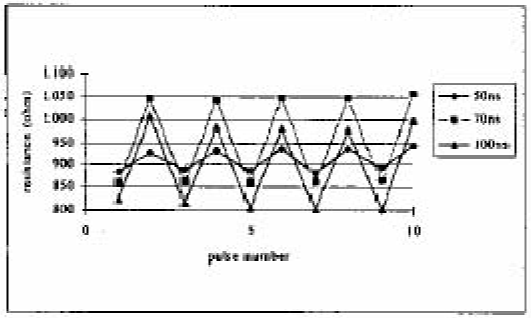
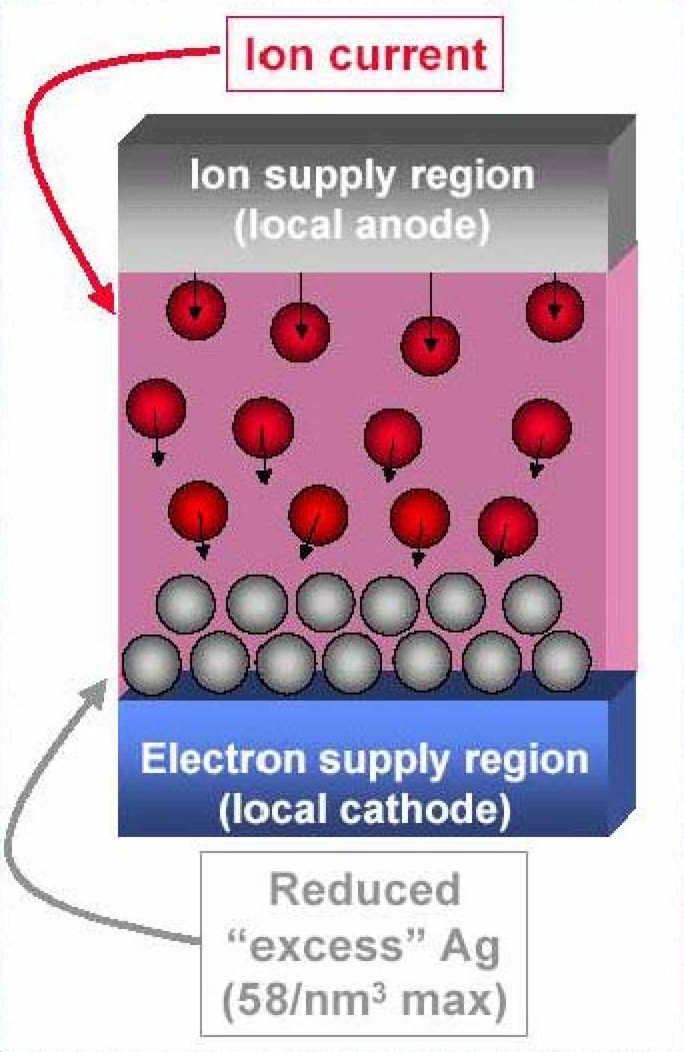


Fig.8 Positive pulse writes the resistor to high resistance-state. Negative pulse reset the resistor to low resistance-state. Programming efficiency depends strongly on pulse width

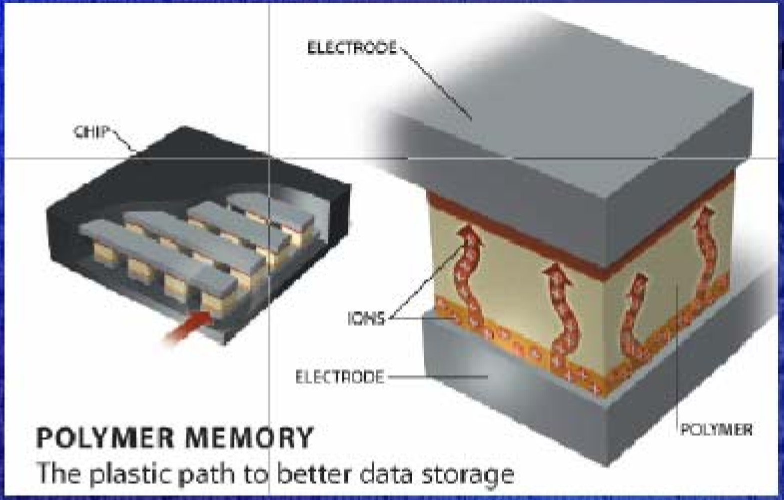
Programmable Metallization Cell

- Operation
  - Silver "dissolved" in chalcogenide
  - Change in resistance with applied electric field driving silver to form low resistance path
  - Reversible with reverse field
  - Low field read with no disturb
  - Very fast write and relatively low power

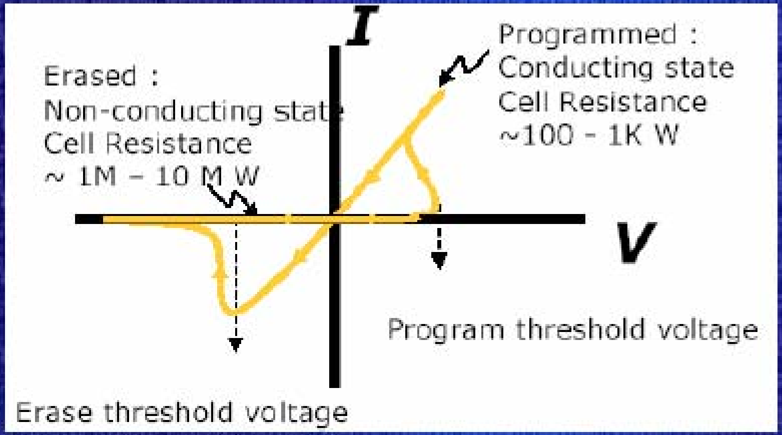


Resistance Polymer Memory

- Operation
  - Polymer material with special formulation
  - Change in resistance due to ionic transport with applied electric field
  - Low resistance when ionic conductance paths formed, high resistance when process is reversed when paths broken
  - Low field read with no read disturb
  - Low cost with polymer

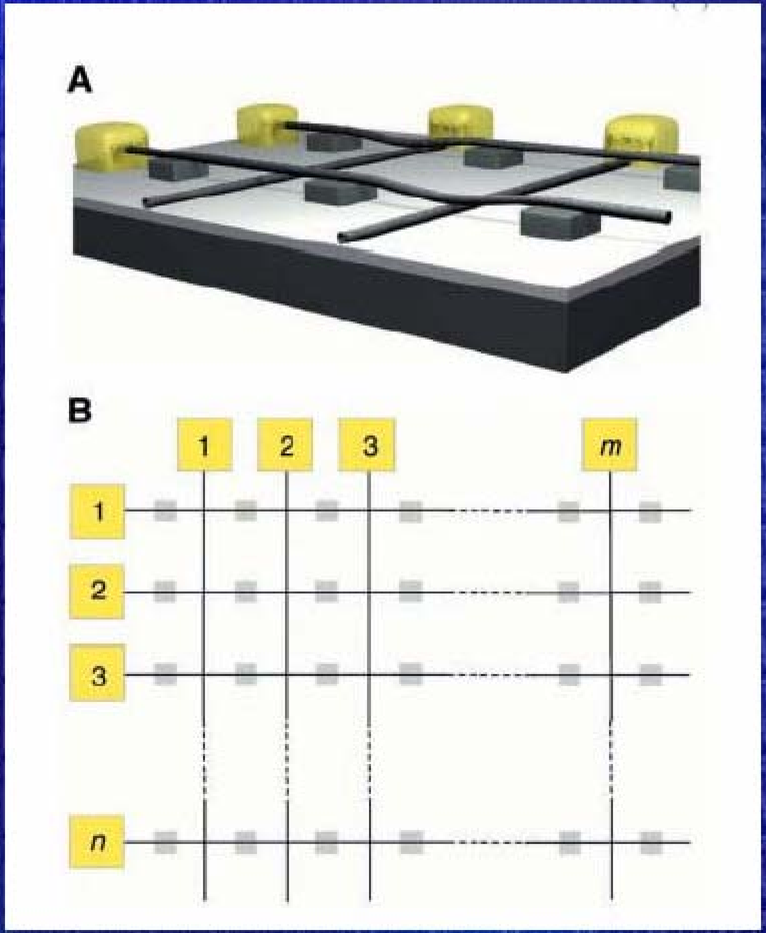


POLYMER MEMORY
The plastic path to better data storage



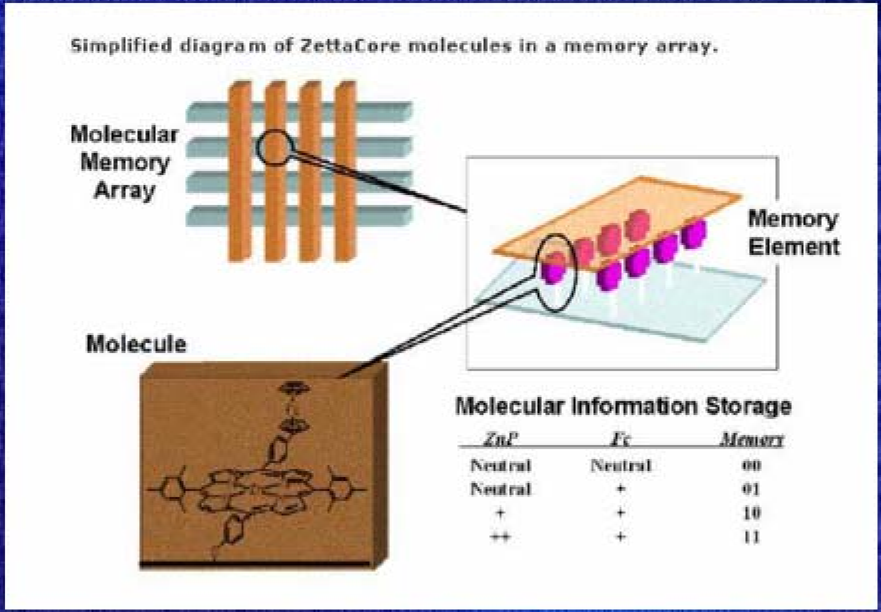
Carbon Nanotube Switches

- Operation
  - Carbon nanotube suspended in crossbar architecture
  - Electrostatic field attracts the nanotube and held together by van der Waals force
  - Reverse bias repulse the wires
  - Reading by low vs high resistance
  - Array Architecture TBD



Molecular Memory

- Operation
  - Molecules in cross point array
  - Change in electronic states in Redox process with applied voltage
  - Memory state sensed by charge displacement
  - More than 1 state can be stored by design of molecule (up to 8 demonstrated)



Simplified diagram of ZettaCore molecules in a memory array.

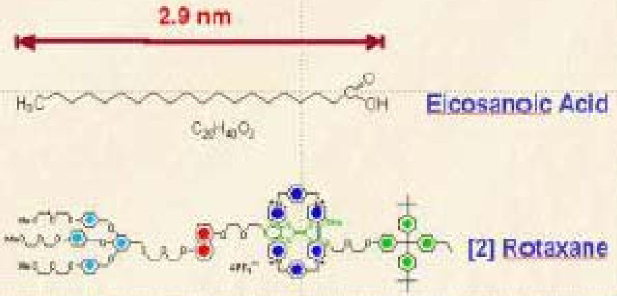
Molecular Information Storage

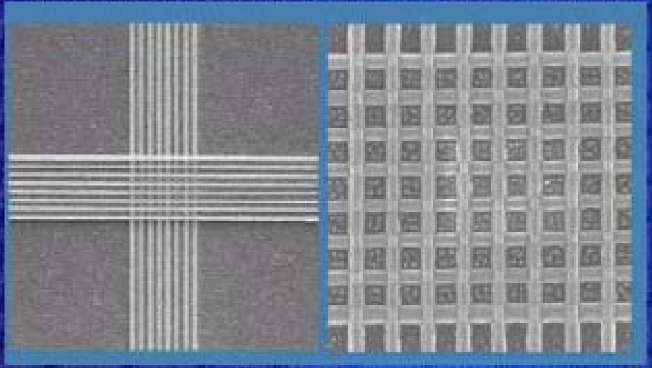
ZnP	Fc	Memory
Neutral	Neutral	00
Neutral	+	01
+	+	10
++	+	11

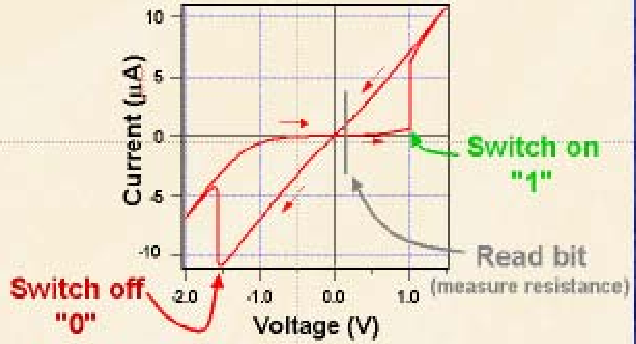
Molecular Tunnel Memory

- Operation
  - Molecules in cross point array
  - Change in resistance with applied electric field
  - Memory state sensed by measured resistance change

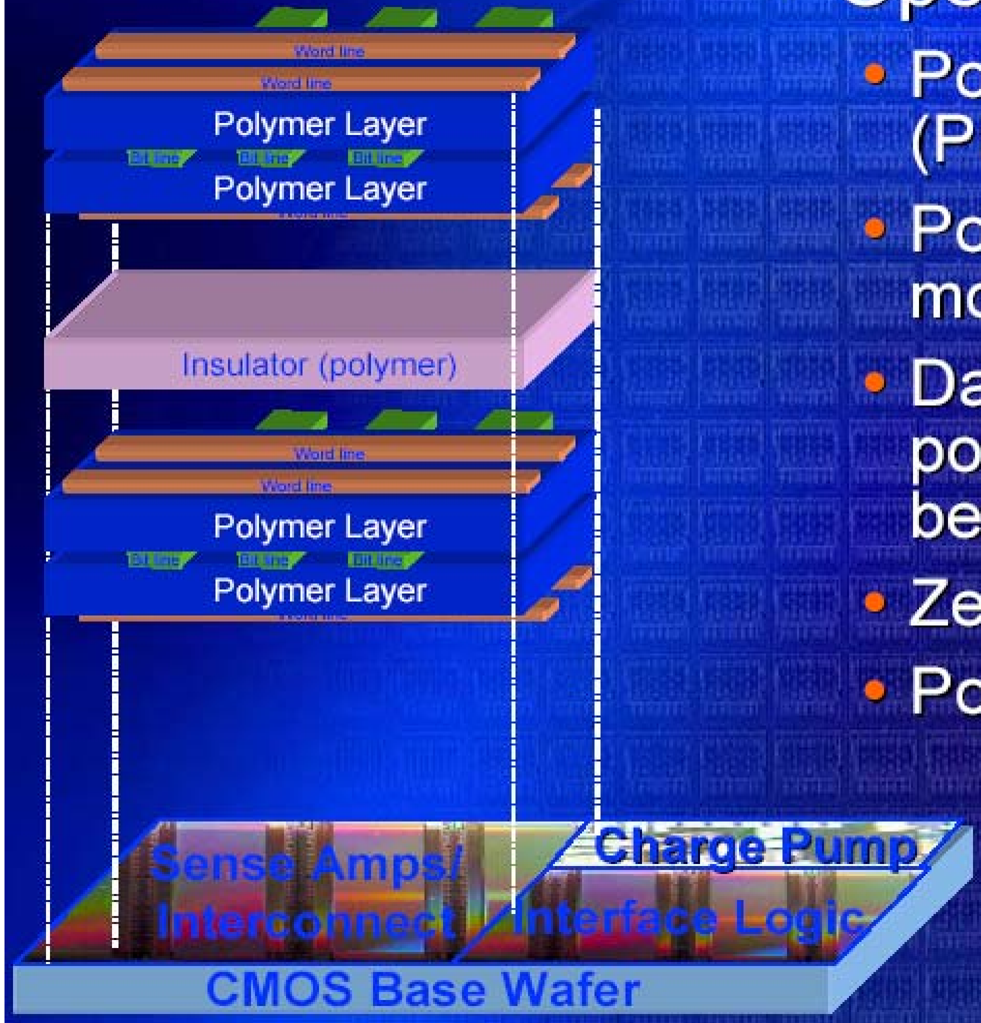
Realization of a Molecular-Tunneling Switch



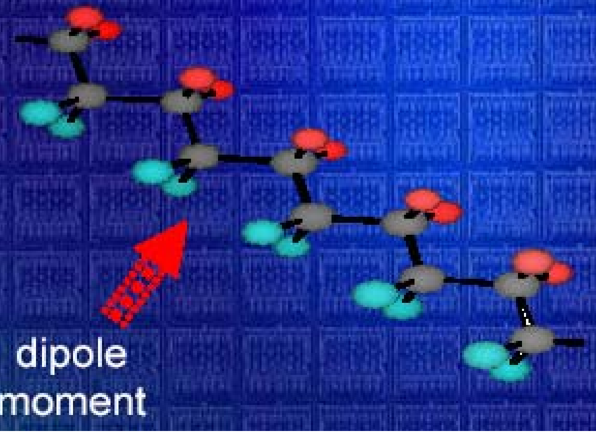




Ferroelectric Polymer Memory

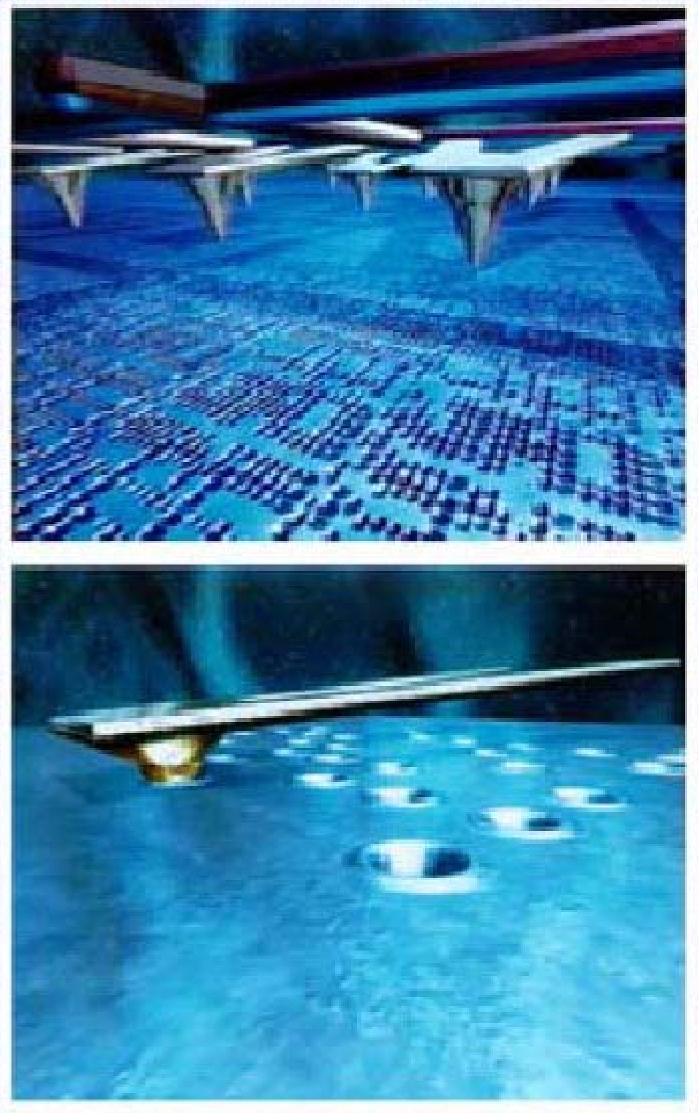


- Operation
  - Polymeric Ferroelectric RAM (PFRAM)
  - Polymer chains with a dipole moment
  - Data stored by changing the polarization of the polymer between metal lines
  - Zero transistors per bit of storage
  - Polymer layers can be stacked



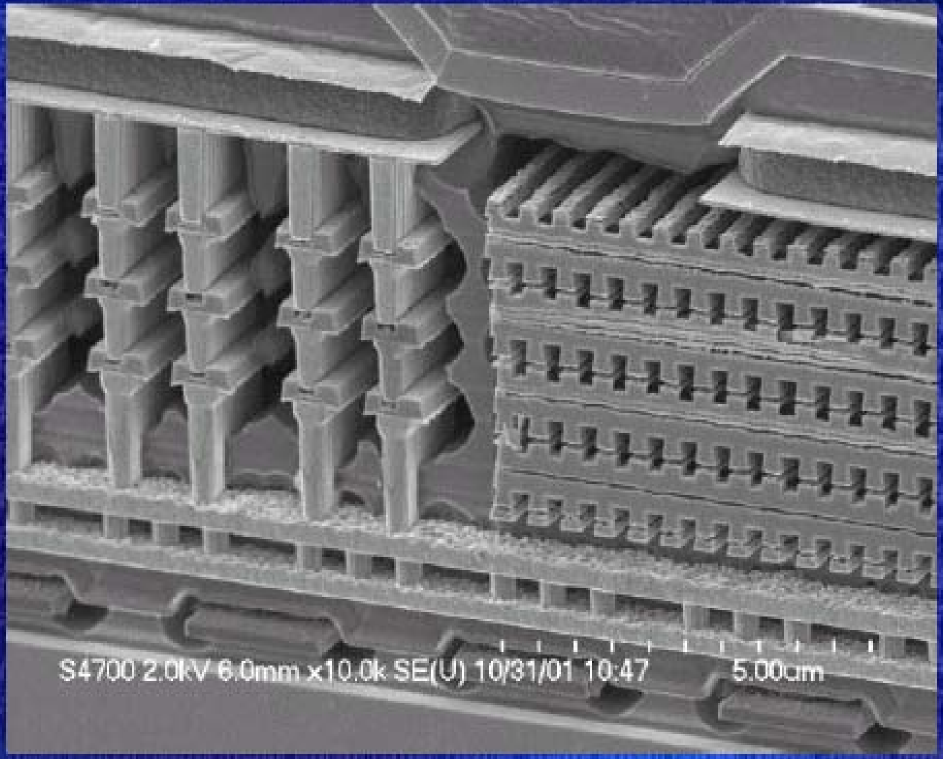
Millipede Memory

- Operation
  - Basic concept similar to punch cards
  - Thermally assisted, rewritable displacement media, PMMA
  - Large array of independently Z-axis controlled tips
  - Low moving mass
  - Direct point to point motion, no rotational latency
  - Low power, no motors, actuators
  - High data transfer rates, concurrent data transfer to/from multiple tips
  - Packaging similar to hermetic devices



3D One Time Program Memory

- Operation
  - Multi-layer one time programmable diodes at minimum lithography dimensions
  - Simplified architecture with minimum number of metal lines
  - Zero transistor per cell
  - Standard CMOS transistors under and outside array, process cost amortized over multiple memory cells
  - Significant lower cost compared to single layer



Spin Polarized E Beam Magnetic Memory

- Operation
  - Basic concept similar to Magnetic Hard Drive
    - Mechanical arm -> spin polarized E Beam
    - Spinning Magnetic Media -> steered E Beam
  - Much lower first read latency, first access limited by E Beam steering; fast read with E Beam
  - Lower power, no motor, no arm actuator
  - No moving part, rugged, vacuum space required
  - Not as small or compact as all solid state memories