Nano-Scale Memory Devices:
Space-Time-Energy Trade-offs

Ralph Cavin and Victor Zhirnov

Semiconductor Research Corporation

Main Points

- Many candidates for beyond-CMOS nano-electronics have been proposed for memory, but no clear successor has been identified.
  - Methodology for system-level analysis
- **How is maximum performance related to device physics?**

Three integrated components of a Memory Device:

- 1) '**Storage node**'
  - physics of memory operation
- 2) **'Sensor'** which reads the state
  - e.g. transistor
- 3) **'Selector'** which allows a memory cell in an array to be addressed
  - transistor
  - diode

- All three components impact scaling limits for all memory devices

Space-Time-Energy Metrics

- Essential parameters of the memory element are:
  - cell size/density,
  - retention time, access time/speed
  - operating voltage/energy.
- None of known memory technologies, perform well across all of these parameters
- At the most basic level, for an arbitrary memory element, there is interdependence between operational voltage, the speed of operation and the retention time.

- More generally, cell dimensions are also part of the trade-off, hence the Space-Time-Energy compromise

Space-Action Principle for Memory

$$Energy \times time \times Volume = \min$$

$$E \cdot t \cdot V = \min$$

$$E \cdot t \cdot L = \min$$

$$E \cdot t \cdot N_{at} = \min$$

The Least Action principle is a fundamental principle in Physics

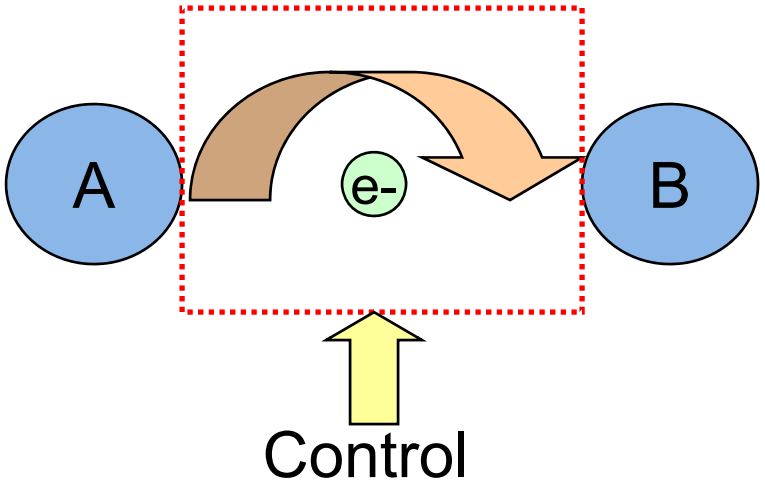
$$E \cdot t = \min \quad (\geq h)$$

Plank's constant
h=6.62x10^{-34} Js

Three Major Memory State Variables

- **Electron Charge ('moving electrons')**
  - e.g. DRAM, Flash
- **Electron Spin ('moving spins')**
  - (STT-) MRAM
- **Massive particle(s) ('moving atoms')**
  - e.g. ReRAM, PCM, Nanomechanical, etc.

Charge-based Memories



DRAM/SRAM

Floating Gate Memory

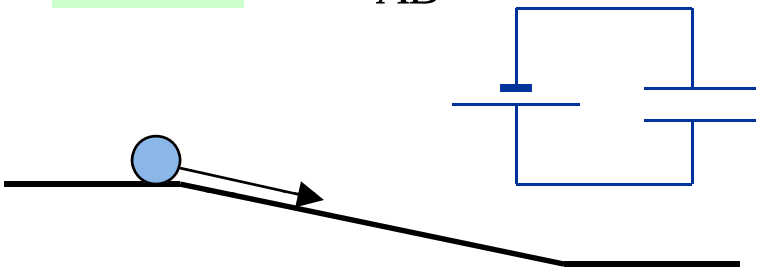
SONOS

Requirements:

1) Efficient charge injection during programming
2) Suppressed back-flow of charge in store/read modes
3) Efficient erase
4) Min. charge/bit: $q=e=1.6x10^{-19}$ Q

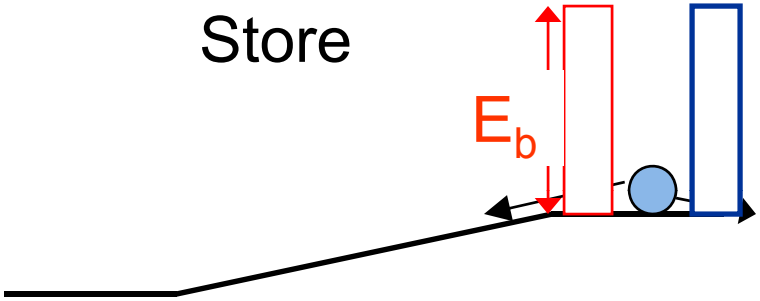
Barrier-less Ohmic Transport: The most efficient injection, but…

Write $I_{AB} \sim \Delta V$



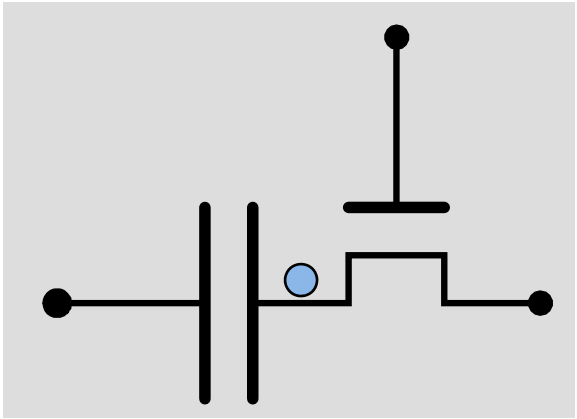
… **difficult retention**

Store



Charge-based memory is a two-barrier system

Example:
DRAM

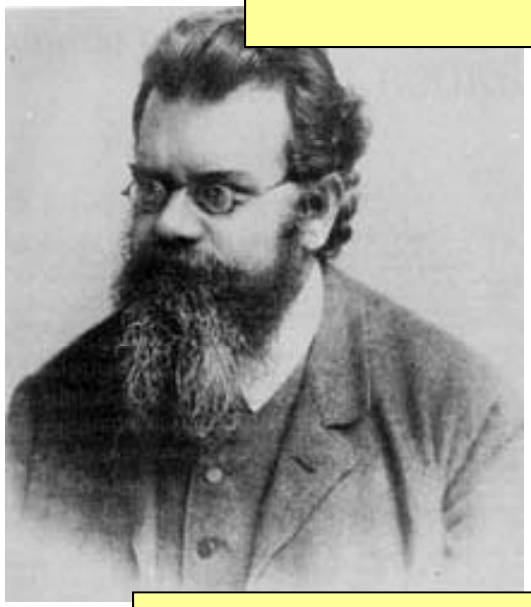


Essential Physics

The operation of charge-based memory devices is governed by these basic equations, which put fundamental constraints on device and circuit parameters.

Boltzmann probability of thermal excitation

$$\Pi = \exp(-\frac{E_b}{k_B T})$$



$$I = I_0 \exp(-\frac{E_b}{k_B T})$$

Heisenberg Relations

$$\Delta x \Delta p \geq \hbar$$

$$\Delta E \Delta t \geq \hbar$$

$$I = I_0 \exp\left(-\frac{2\sqrt{2m}}{\hbar}(a\sqrt{E_b})\right)$$

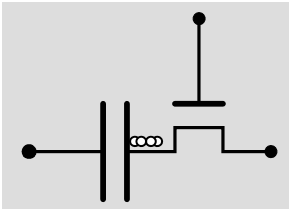
What is the minimum barrier height for the charge-based memory?

Thermionic leakage current (ideal case):

$$J_{th} = J_{th0} \cdot \exp\left(-\frac{E_b}{kT}\right)$$

$$q = CV$$

25fF×0.5V×=3.75 fCl



E_b, eV	Max. retention
0.7	4 ms
0.76	24 ms
0.8	77 ms
1.4	1 month
1.57	**12 years**

Standard DRAM requirement: 64 ms

High-barrier are needed for Non-volatile memory

→ E_{bmin} Min. barrier height for NVM

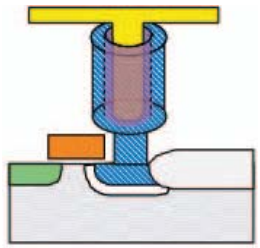
Problem: In Si devices $E_{bmax} < E_g = 1.1$ eV

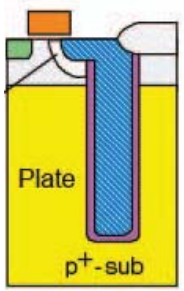
Volatile electron-based memory: DRAM

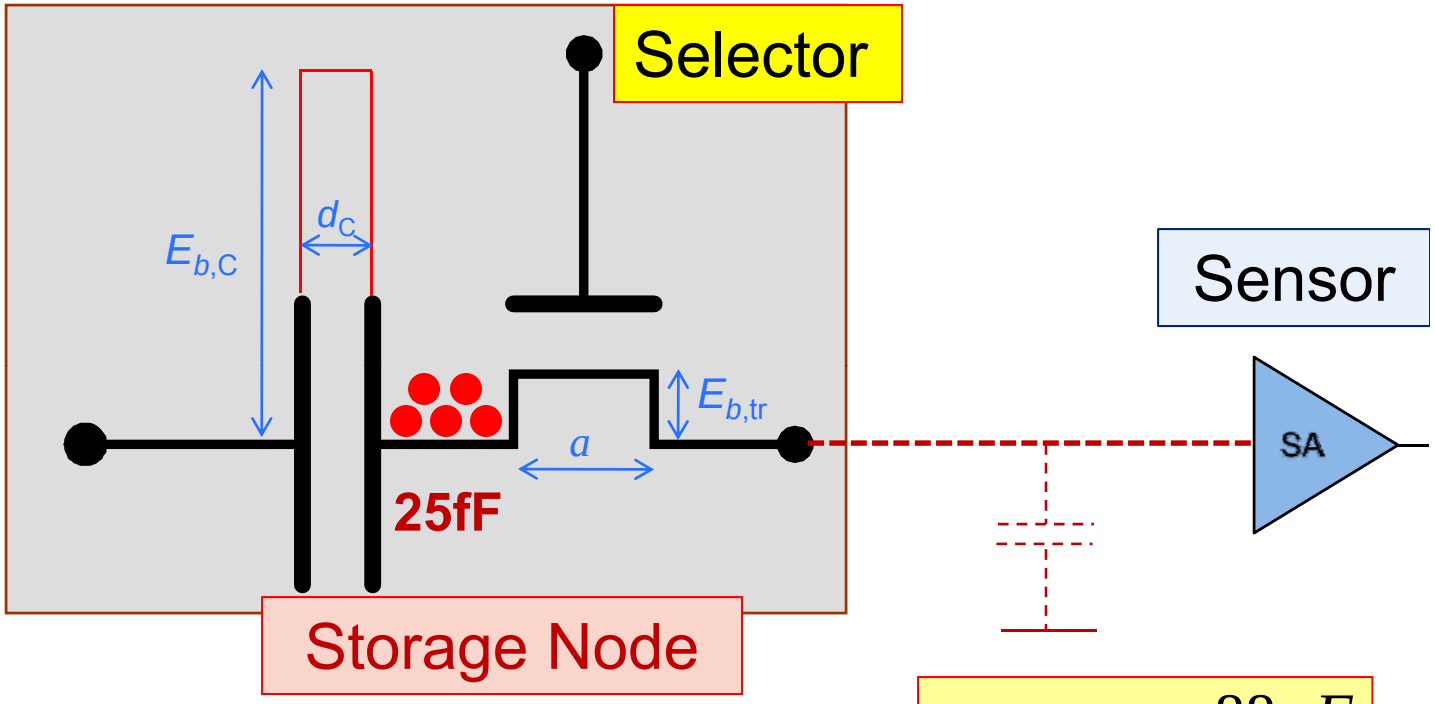
$$N_{el} = \frac{CV}{e}$$

$$E = \frac{CV^2}{2}$$

$$t = \frac{CV}{I}$$







$$C_{\text{int}} \propto \frac{\varepsilon_0}{L} \sim \frac{88\,pF}{1cm}$$

N_{el}~10^5

a=10 nm

V_{cap}=6x10^{-15} cm^3

E~10^{-14} J

t_w~1 ns

DRAM summary

a=10 nm

N_{el}~10^5

E~10^{-14} J

V_{cap}=6x10^{-15} cm^3

t~1 ns

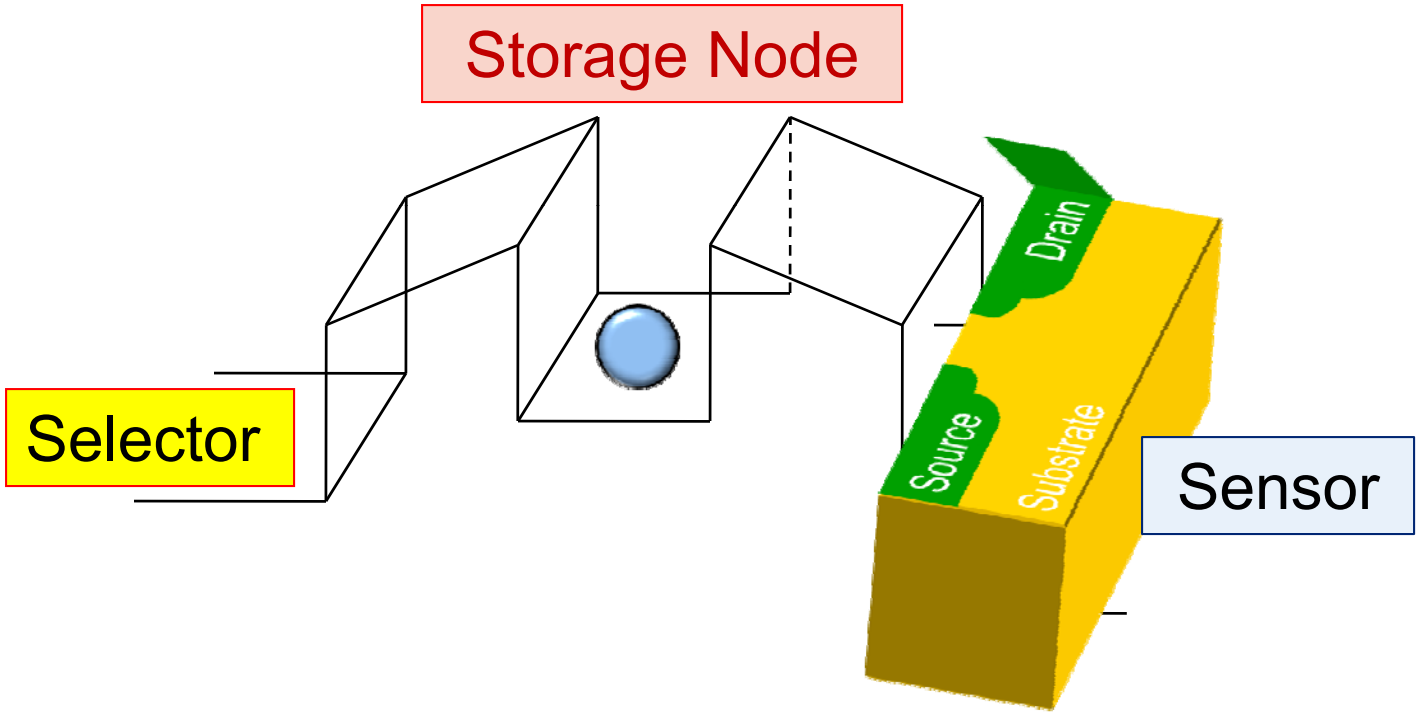
E×t×V~10^{-9} J-ns-nm^3

DRAM inherent issues:

Selector -Low barrier height-***Volatility***

Sensor - Remote sensing – ***Large size of Storage node***

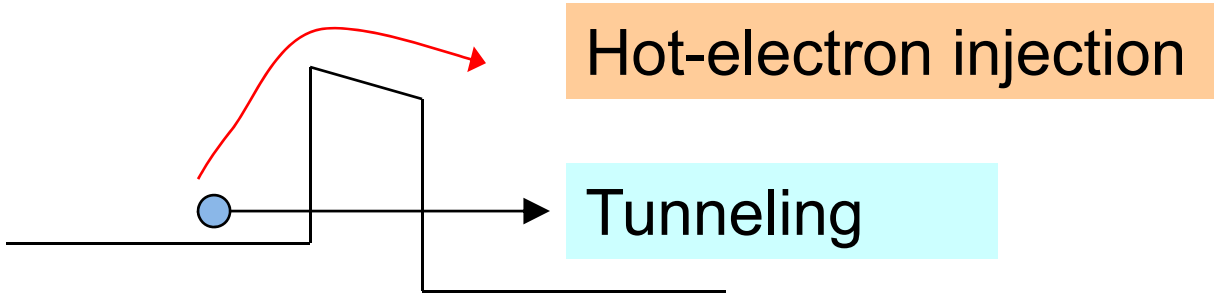
Flash: Local Sensing of Memory State



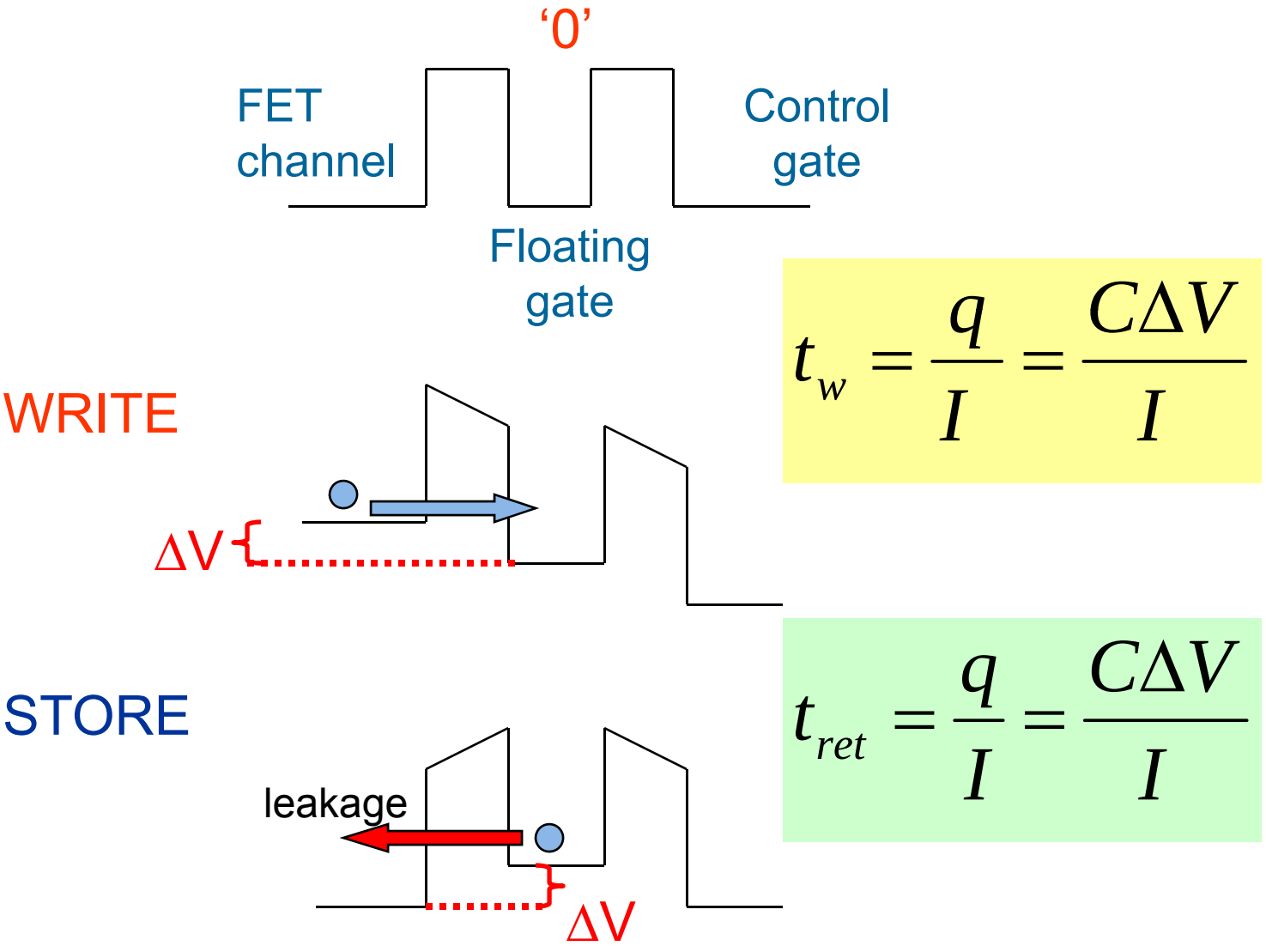
Charge injection problem in high-barrier systems

High-barriers are needed for Non-volatile memory

BUT: Barrier formed by an insulating material (large E_b) cannot be suppressed) – charge transport in the presence of barriers: ***Non-ohmic charge transport***



Floating gate memory: WRITE and STORE modes



$$t_w = \frac{q}{I} = \frac{C\Delta V}{I}$$

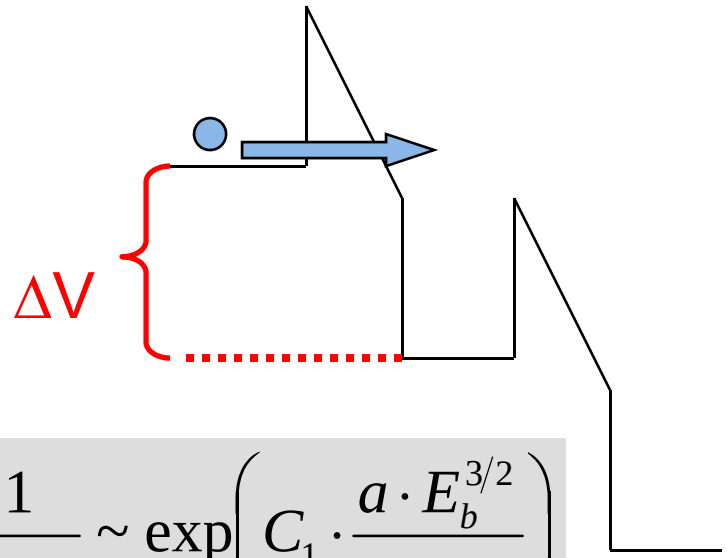
$$t_{ret} = \frac{q}{I} = \frac{C\Delta V}{I}$$

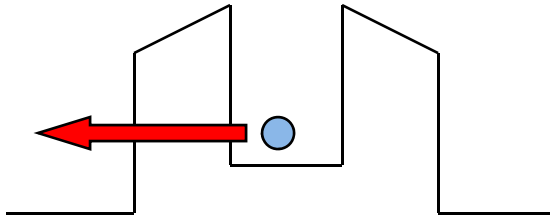
We need to create an asymmetry in charge transport through the gate dielectric to maximize the I_{write}/I_{ret} ratio

Floating gate cell:

Write – **triangle barrier**

Retention – **trapezoidal barrier**





$$t \sim \frac{1}{T_{F-N}} \sim \exp\left(C_1 \cdot \frac{a \cdot E_b^{3/2}}{\Delta V}\right)$$

$$t_{ret} \sim \frac{1}{T_{DT}} \sim \exp\left(C_2 \cdot a \cdot \left\langle \sqrt{E_b} \right\rangle\right)$$

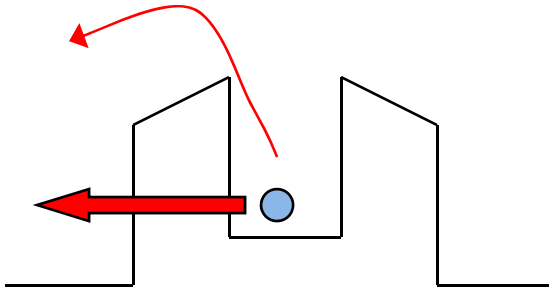
The asymmetry in charge transport between WRITE and STORE modes is achieved through different shape of barrier (triangle vs. trapezoidal)

Retention Analysis:
Minimum Barrier Height and Width:

$$t_s = \frac{Ne}{I_s}$$ ~10 y

$$I_{th} = I_{th0} \cdot \exp\left(-\frac{E_b}{kT}\right)$$

E_{bmin}=1.8 eV (400K)



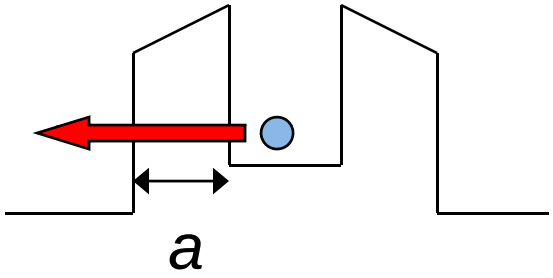
$$I_{tun} = A\frac{e^2}{h^2}\frac{\sqrt{2mE_b}}{a}V\exp\left(\frac{2\sqrt{2m}}{\hbar}\cdot a\cdot\sqrt{E_b - \frac{eV}{2}}\right)$$

a_{min}=6.9 nm

Floating Gate Cell Retention and WRITE characteristics

Retention: direct tunneling

$V_{stored} < E_b$

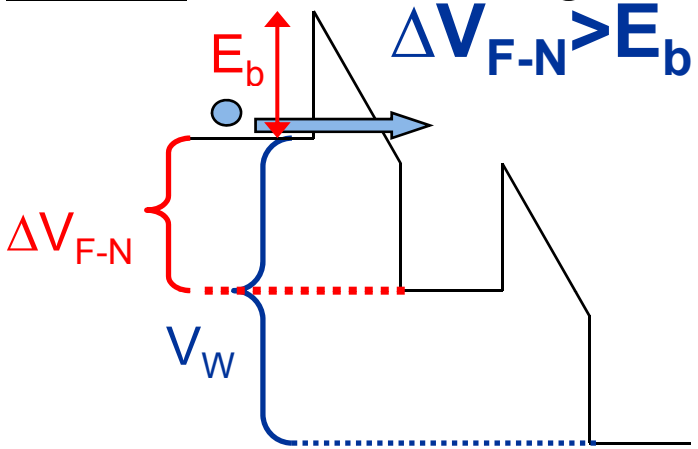


The retention time strongly depends on thickness

Ideal case

Barrier	Eb	V Ret	a	t Ret
Si/SiO2	3.1 eV	2 V	4 nm	4.35 min
Si/SiO2	3.1 eV	2V	5.4 nm	20 y
Min. barrier	**1.8 eV**	**0.9 V**	**6.9 nm**	**11 y**

Write: F-N tunneling:



$$V_{W\min} = 2 \cdot \Delta V_{F-N} \geq 2E_b$$

Si/SiO_2: E_b=3.1 V, V_{WSiO2}>6.2 V

For lower WRITE voltage E_b should be decreased:

E_{bmin}=1.8 eV V_{min}>4 V

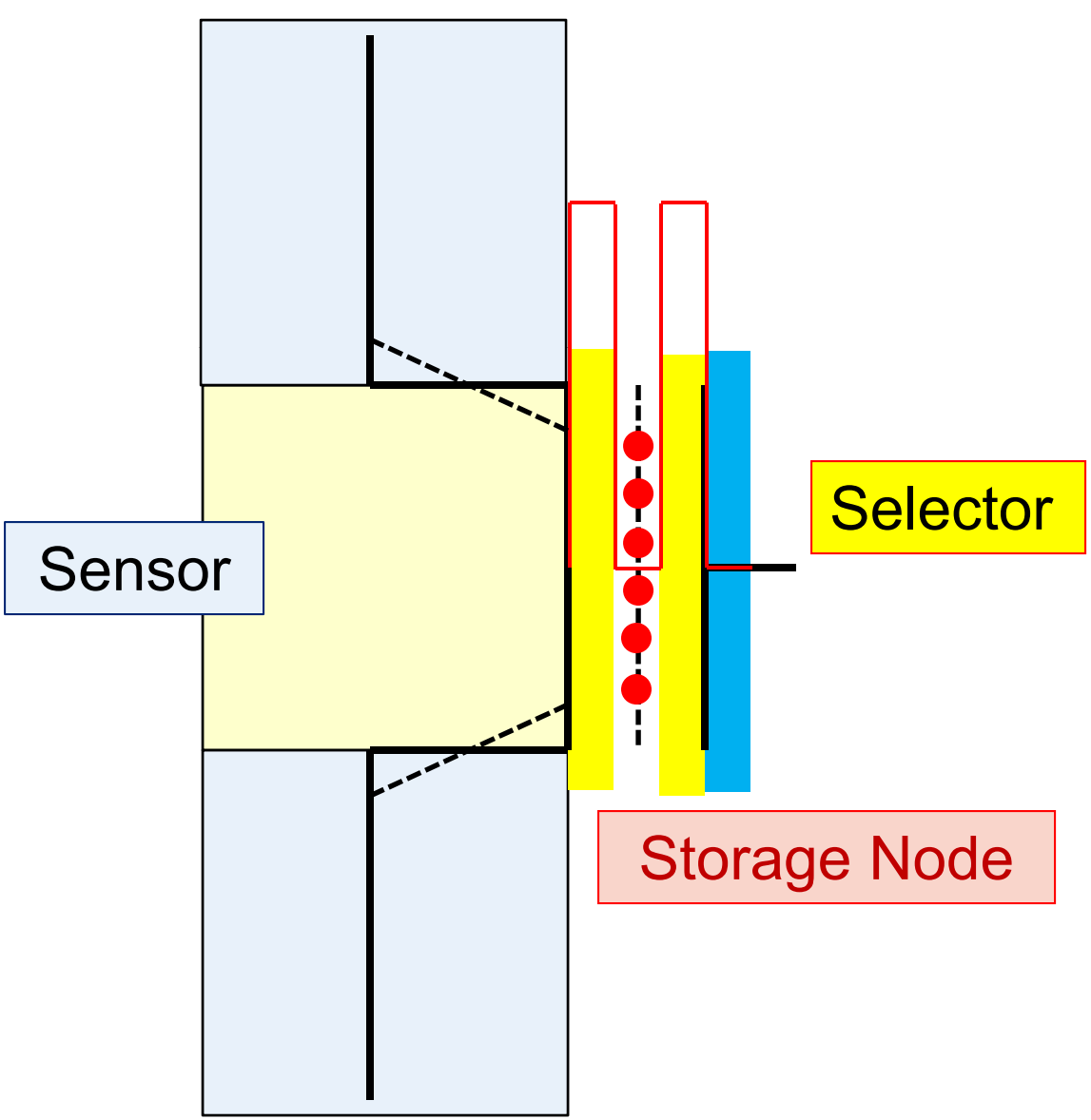
Barrier	Eb	V WR	a	t WR
Si/SiO2	3.1 eV	6.8 V	5.4 nm	1h
Si/SiO2	3.1 eV	12 V	5.4 nm	30ms
Min. barrier	**1.5 eV**	**6 V**	**6.9 nm**	**40ms**

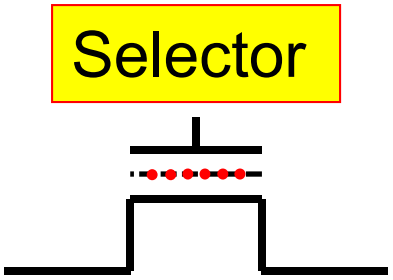
Voltage-Time Dilemma

- For an arbitrary electron-charge based memory element, there is interdependence between operational voltage, the speed of operation and the retention time.

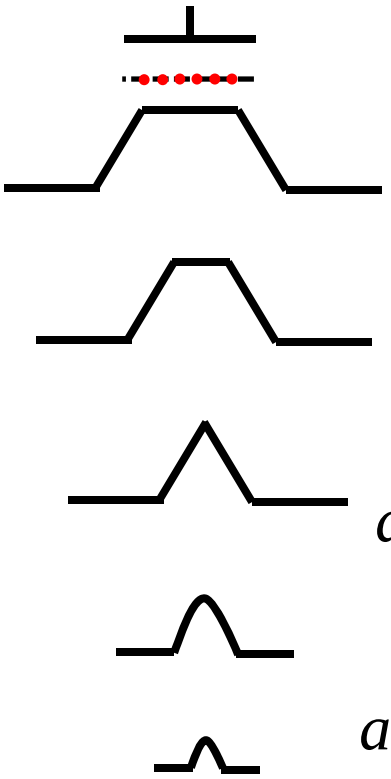
- Specifically, the nonvolatile electron-based memory, suffers from the “barrier” issue:
  - High barriers needed for long retention do not allow fast charge injection
  - It is difficult (impossible?) to match their speed and voltages to logic

Flash Scaling limits due to ‘Sensor’



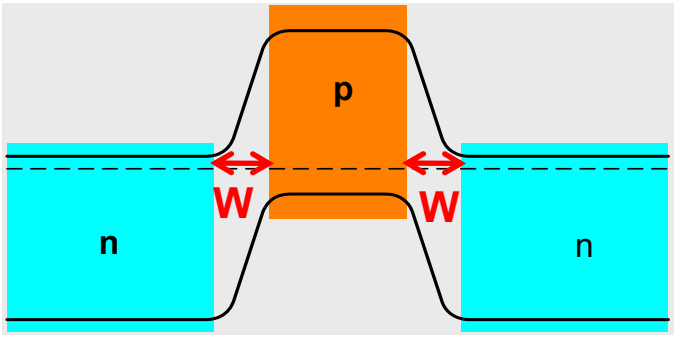


Semiconductor Physics sets limits on barrier quality



$a=2W$

$a<2W$



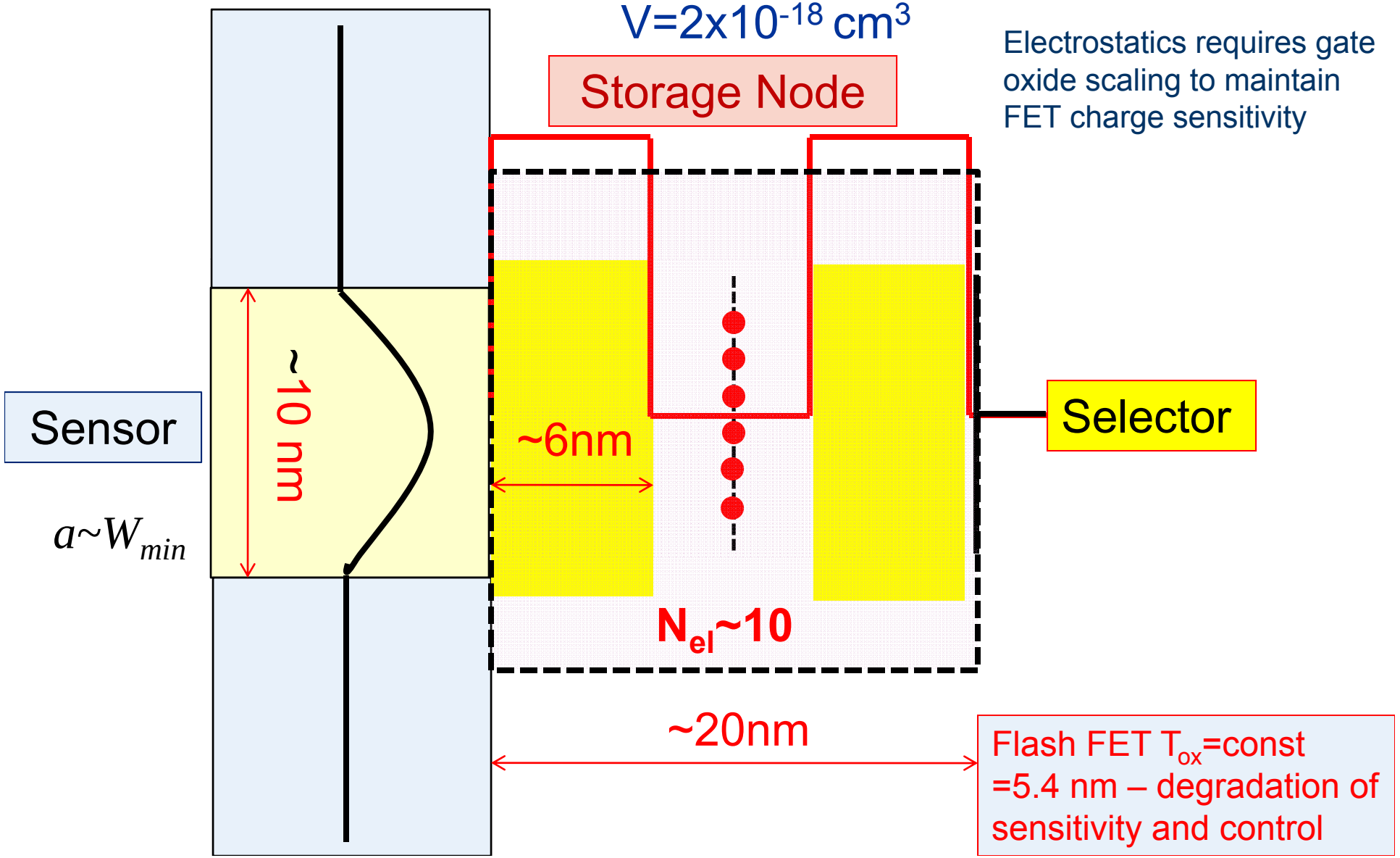
Depletion width

$$W \approx \sqrt{2\varepsilon\varepsilon_0 V_{bi} \frac{N_A + N_D}{N_A N_D}}$$

$$W_{\min} = \sqrt{2\varepsilon\varepsilon_0 E_g \frac{N_C + N_V}{N_C N_V}} \sim 10\,nm$$

N_C – effective density of states in the conduction band, for Si N_C=2.8x10^{19} cm^{-3}
N_V - effective density of states in the valence band, for Si N_V=1.4x10^{19} cm^{-3}
E_g – the band gap, for Si E_g=1.12 eV

Flash in the limits of scaling



Flash Summary

a=10 nm

N_{el}~10

E~10^{-15} J

t~1 μs=1000 ns

V=2000 nm^3

E×t×V~10^{-9} J-ns-nm^3

Conclusion on ultimate charge-based memories

- All charge-based memories suffer from the “barrier” issue:
  - High barriers needed for long retention do not allow fast charge injection
  - It is difficult (impossible?) to match their speed and voltages to logic
    - **Voltage-Time Dilemma**

Non-charge-based NVMs?

Emerging Memory Devices

The Choice of Information Carrier

Desired: 'Benchmark' memory cell

Driver: Cell Scaling

Cell size, l	<10 nm
Store time, t_s	$>10^8$ s
Write time, t_w	$<10^{-7}$ s
Read time, t_r	$<10^{-7}$ s
Read Voltage, V_r	~1 V
Read current, I_r	~10^{-6} A
Read current density, J_r	$>10^6$ A/cm^2

Min. sense amplifier requirement (R. Waser et al.)

Driver: Sensing

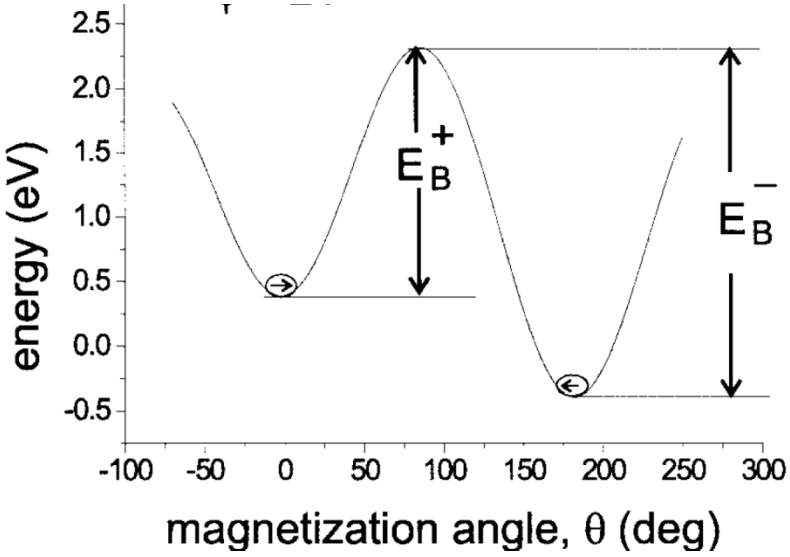
Spin torque transfer MRAM

Magnetic storage node (Moving spins): Energy Limit

$$f_{tr} = f_0 \exp\left(-\frac{E_b}{k_B T}\right) = f_0 \exp\left[-\frac{KV}{k_B T}\right]$$

(f_0~10^9-10^{10} c^{-1})

$$t_{store} = \frac{1}{f_0} \exp\left[-\frac{KV}{k_B T}\right]$$



t_{store}>10 y

the anisotropy constant of a material

$E_b = KV > 36k_BT$~1.25 eV

volume

D. Weller and A. Moser, "Thermal Effect Limits in Ultrahigh-Density Magnetic Recording", IEEE Trans. Magn. 36 (1999) 4423

Magnetic storage node (Moving spins):
Size Limit

$E_b = KV > 36k_BT \sim 1.25$ eV

the anisotropy constant of a material

$K \sim 0.1\text{-}1$ J/cm^3

volume

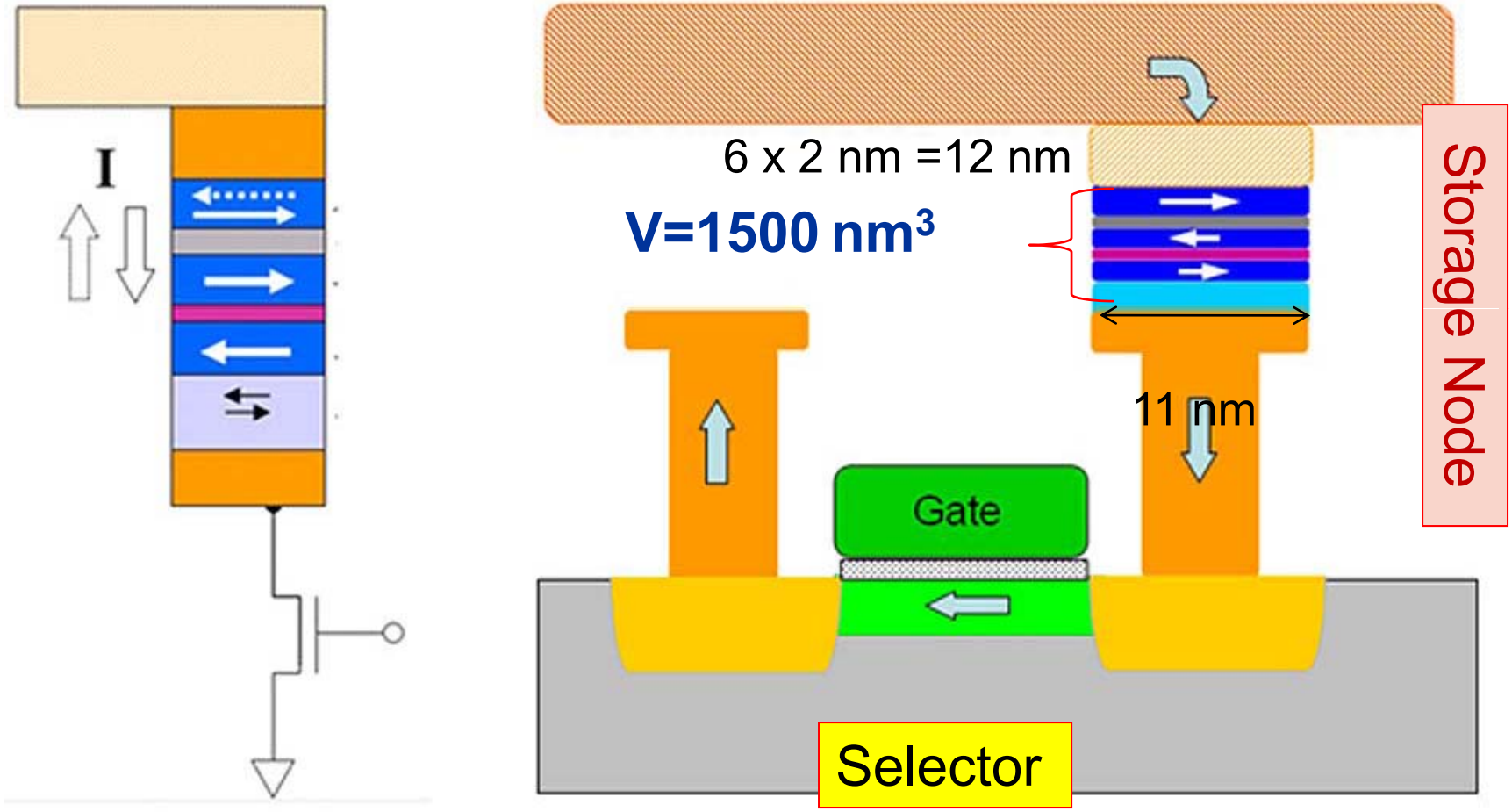
$V = L^2T \sim 2 \times 10^{-19}$ cm^3

Thin film: T=2nm L~11nm

$N_{at} \sim 10^4$

$N_{spin} \sim 10^5$

FET selector is biggest part of STT-MRAM in the limits of scaling



J-G. Zhu, Proc. IEEE 96 (2008) 1786

STT-RAM summary

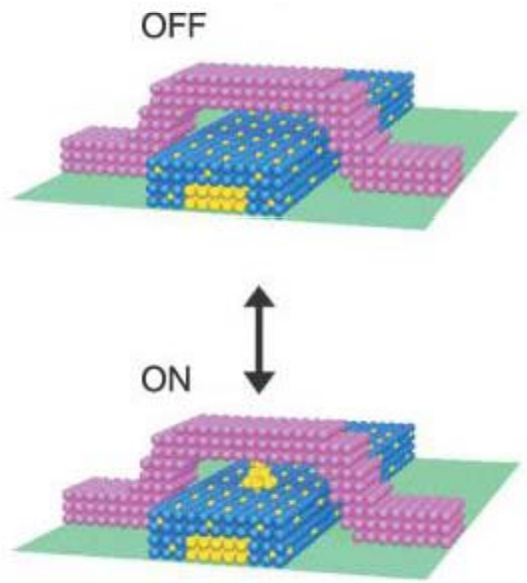
$V=1500\ nm^3$ **$t_w \sim 1\ ns$**

$E \sim 10^7\ A/cm^2 \times 11nm^2 \times 1V \times 1\ ns \sim 10^{-14}\ J$

$E \times t \times V \sim 10^{-11}\ J\text{-}ns\text{-}nm^3$

Scaled ReRAM

Moving atoms: ‘Atomic Relay’



Atomic-scale switch, which opens or closes an electrical circuit by the controlled reconfiguration of silver atoms within an atomic-scale junction.

Such ‘atomic relays’ operate at room temperature and the only movable part of the switch are the contacting atoms, which open and close a nm-scale gap.

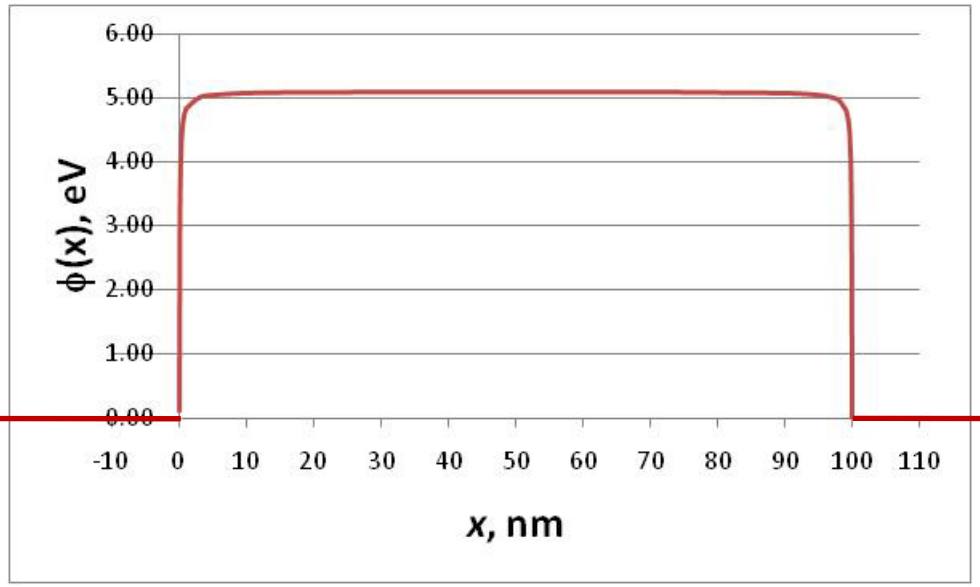
- Small (~1 nm)
- Fast (~1 ns) - projection
- Low voltage (<1V)

Nature **433**, 47-50 (6 January 2005)
Quantized conductance atomic switch
K. Terabe, T. Hasegawa, T. Nakayama and M. Aono

Two-sided barrier

a=100 nm

Au
ϕ_0=5.1 eV



Au
ϕ_0=5.1 eV

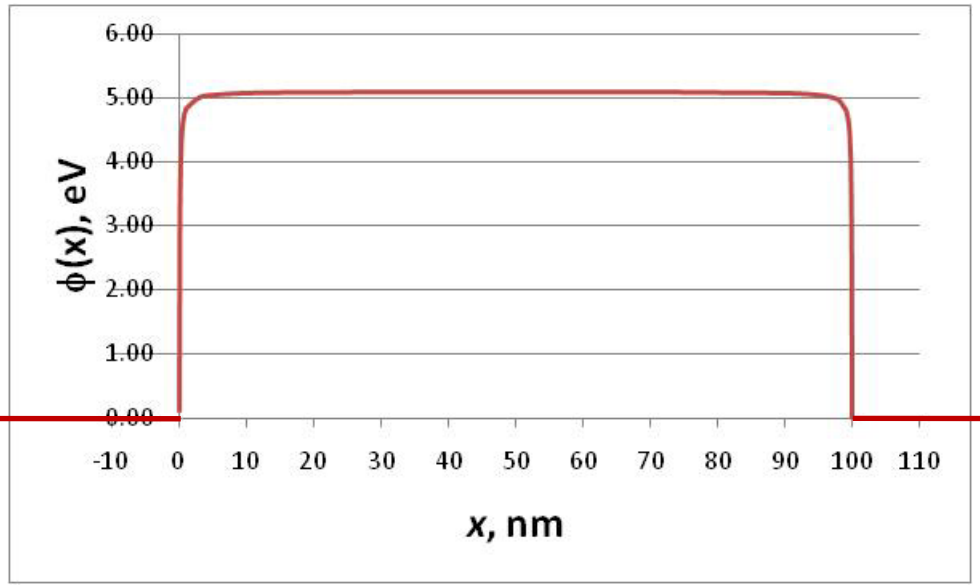
Both electrodes influence the potential of the electron within the electrode separation. For small gaps, the near electrode electric fields will influence the energy barrier

Interface-to-interface interaction

Two-side barrier

a=100 nm

Au
ϕ_0=5.1 eV

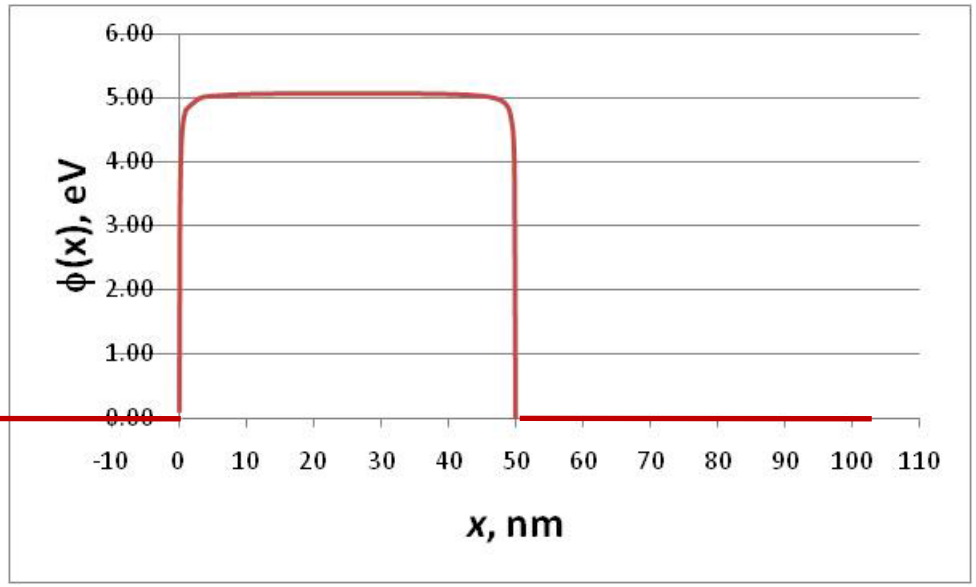


Au
ϕ_0=5.1 eV

Two-side barrier

a=50 nm

Au
ϕ_0=5.1 eV

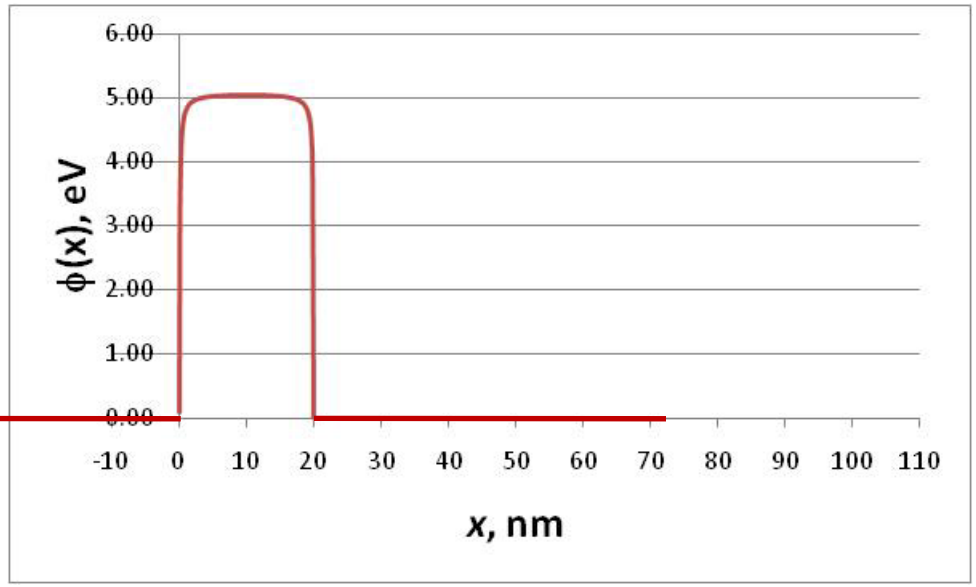


Au
ϕ_0=5.1 eV

Two-side barrier

a=20 nm

Au
ϕ_0=5.1 eV

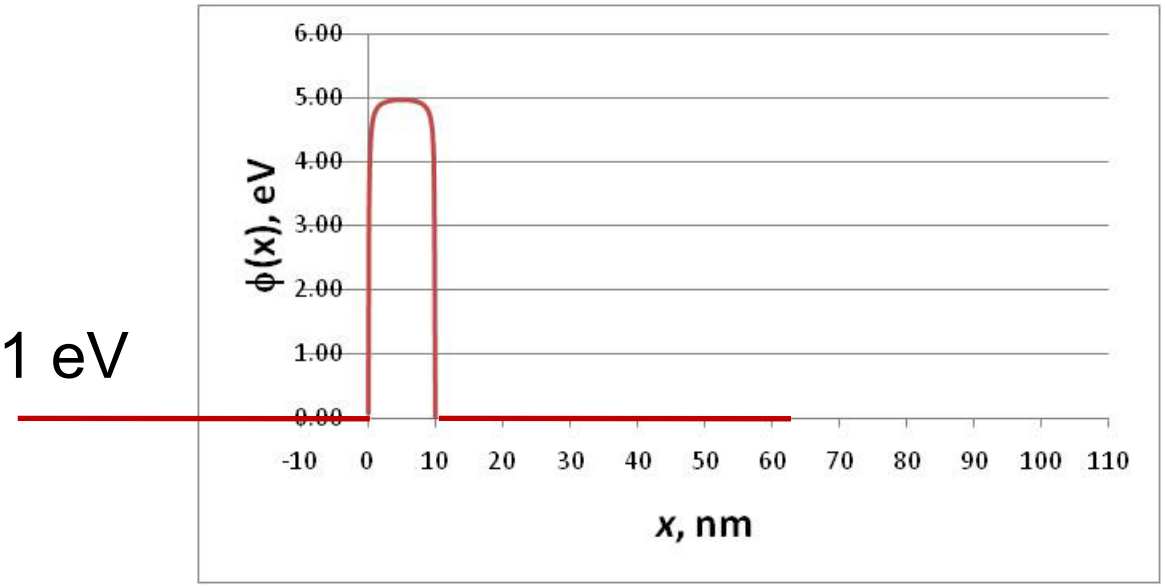


Au
ϕ_0=5.1 eV

Two-side barrier

a=10 nm

Au
ϕ_0=5.1 eV

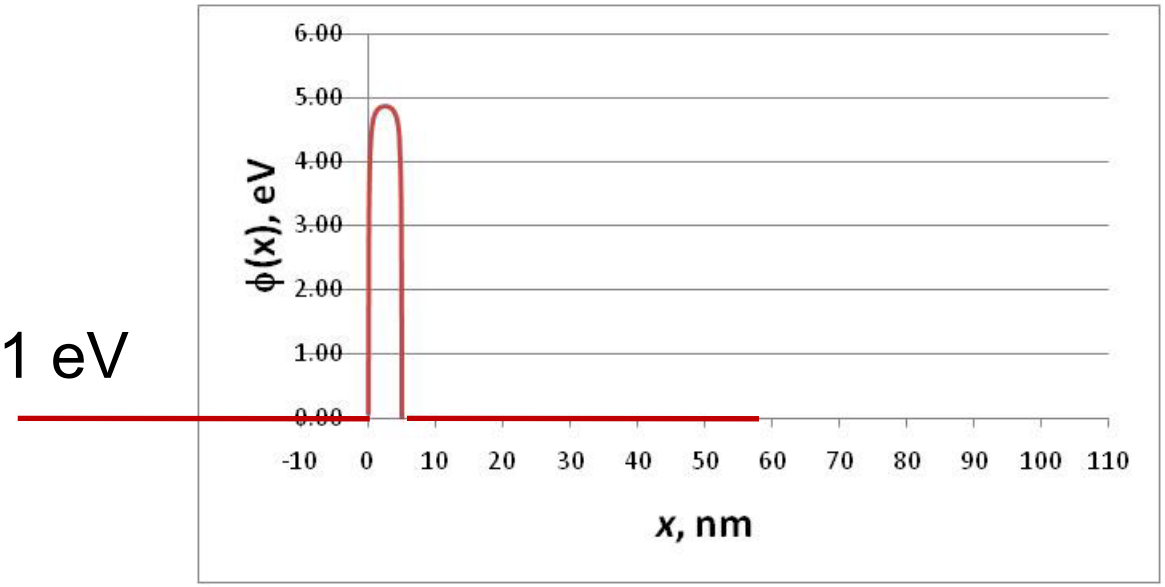


Au
ϕ_0=5.1 eV

Two-side barrier

a=5 nm

Au
ϕ_0=5.1 eV

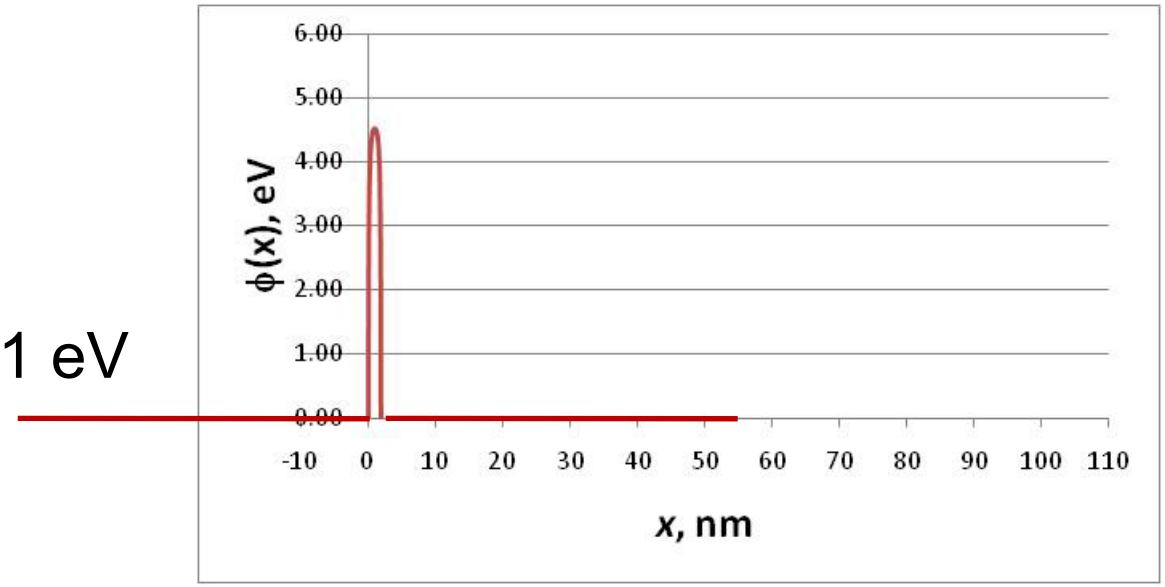


Au
ϕ_0=5.1 eV

Two-side barrier

a=2 nm

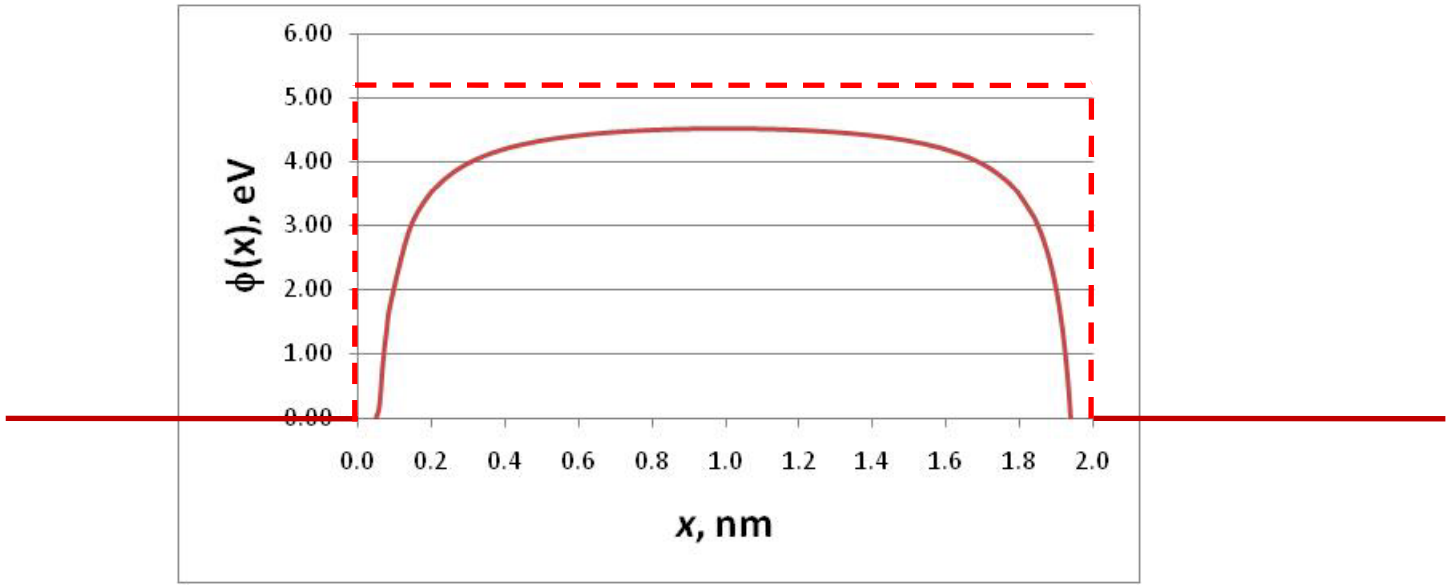
Au
ϕ_0=5.1 eV



Au
ϕ_0=5.1 eV

Two-side barrier

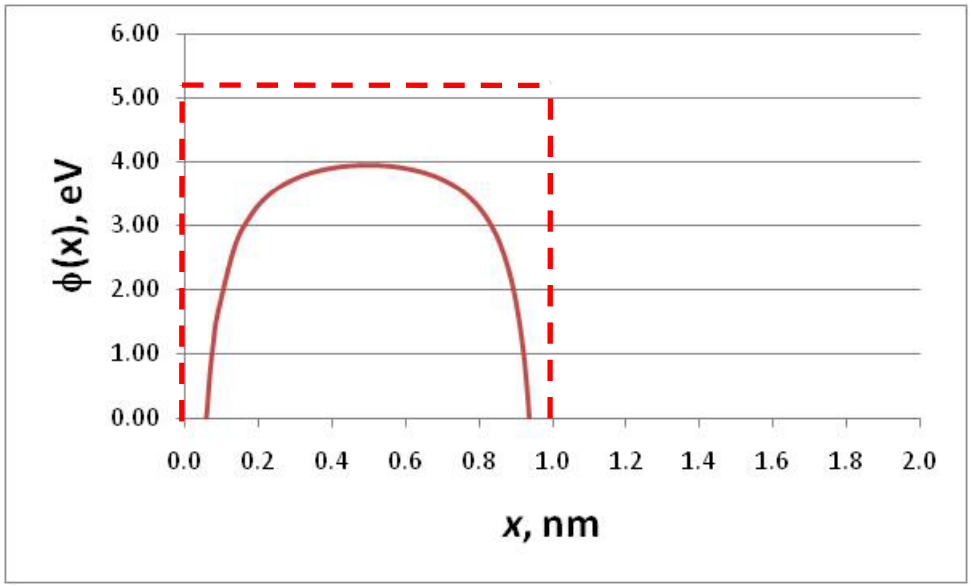
a=2 nm



V=0

Two-side barrier

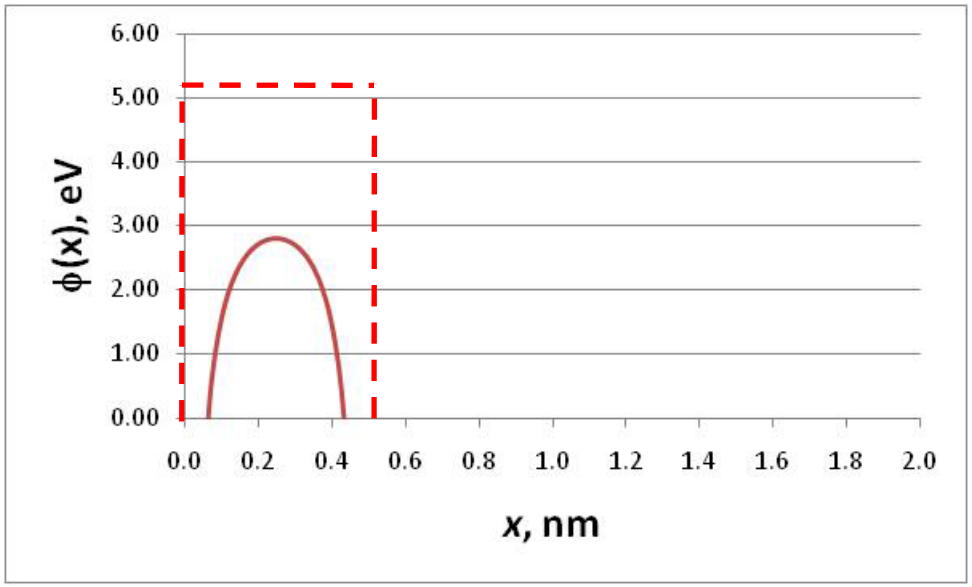
a=1 nm



V=0

Two-side barrier

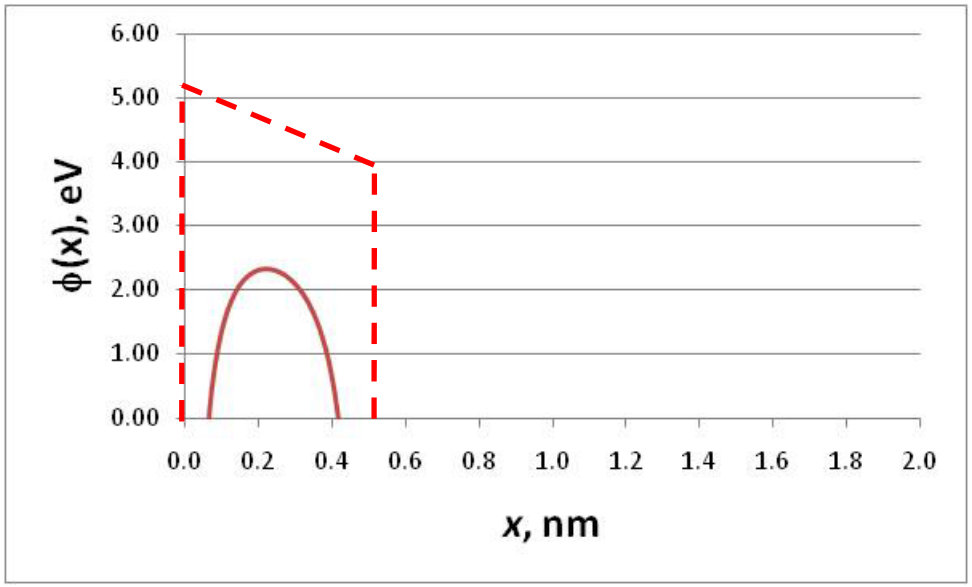
a=0.5 nm



V=0

Two-side barrier

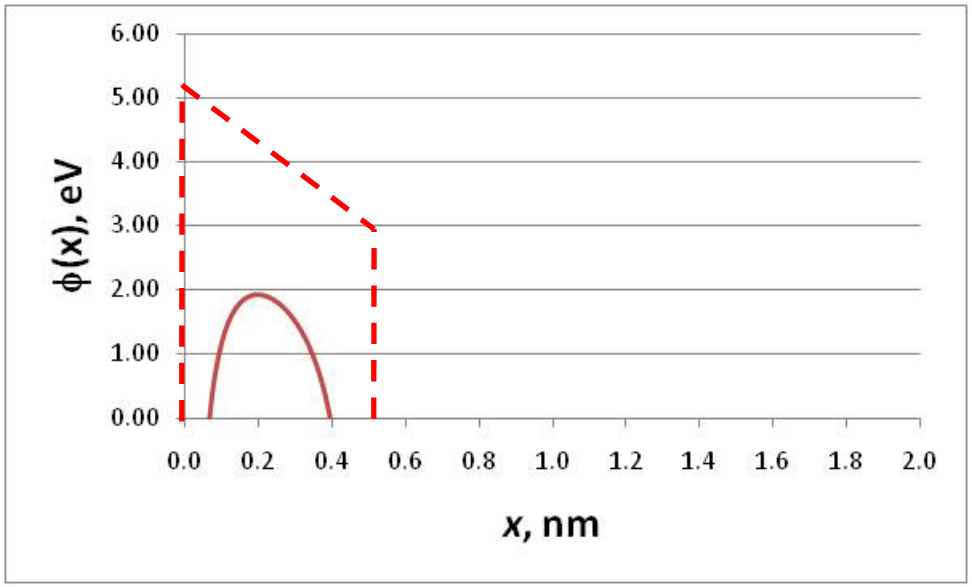
a=0.5 nm



V=1 volt

Two-side barrier

a=2 nm

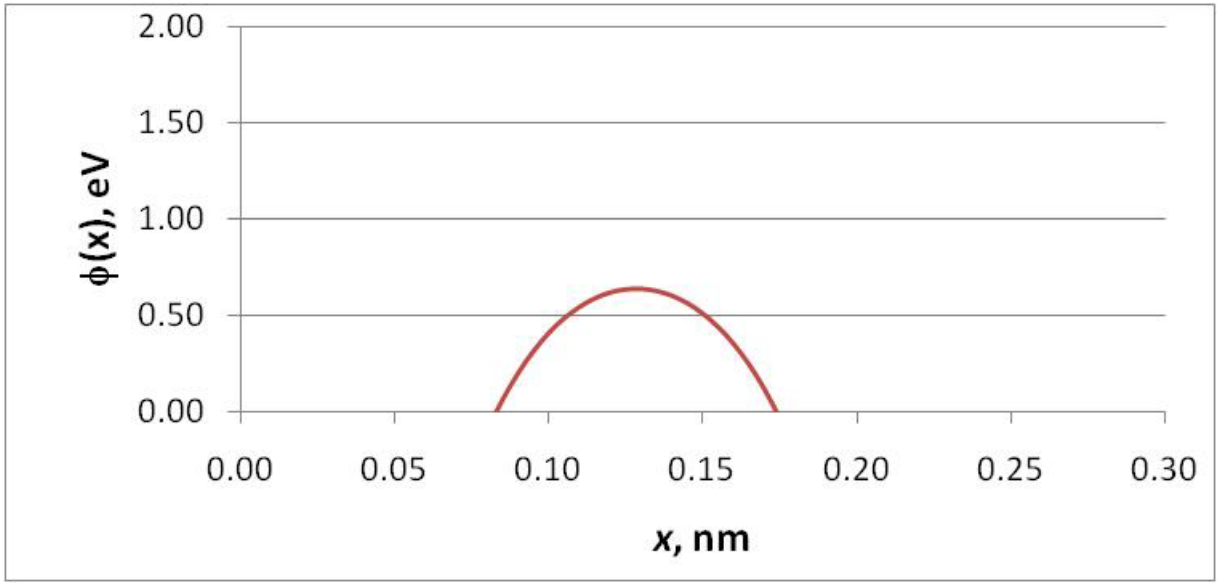


V=2 volt

Ultimate ReRAM: 1-atom gap

$$a \sim n_{Au}^{-\frac{1}{3}} = 0.257nm$$

V=0



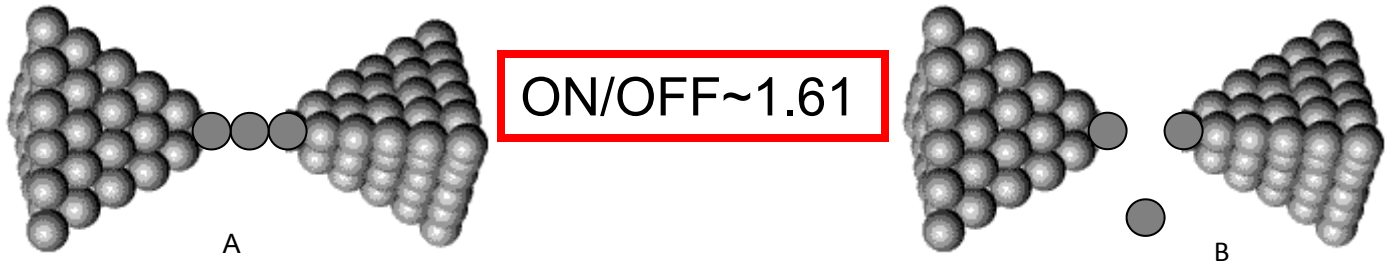
E_b=0.64 eV

d_t=0.09 nm

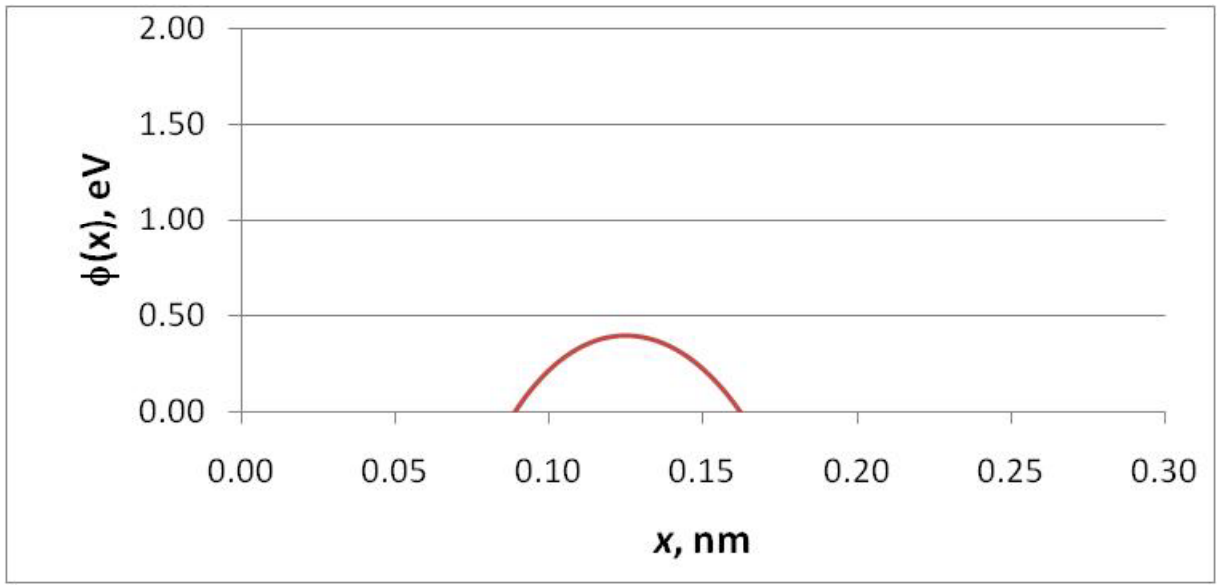
$d_t << a$

Ultimate ReRAM: 1-atom gap

$$a \sim n_{Au}^{-\frac{1}{3}} = 0.257nm$$



V=0.5 V

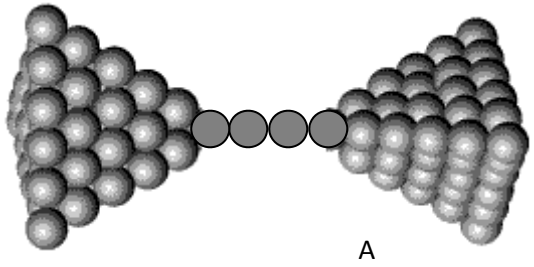


E_b=0.38 eV

d_t=0.075 nm

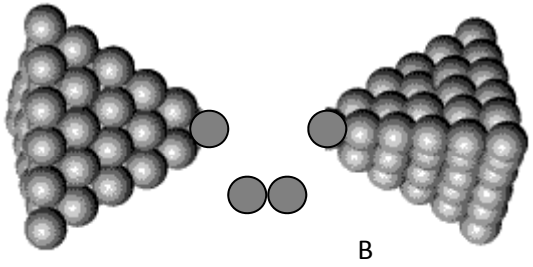
$d_t << a$

Ultimate ReRAM: 2-atom gap

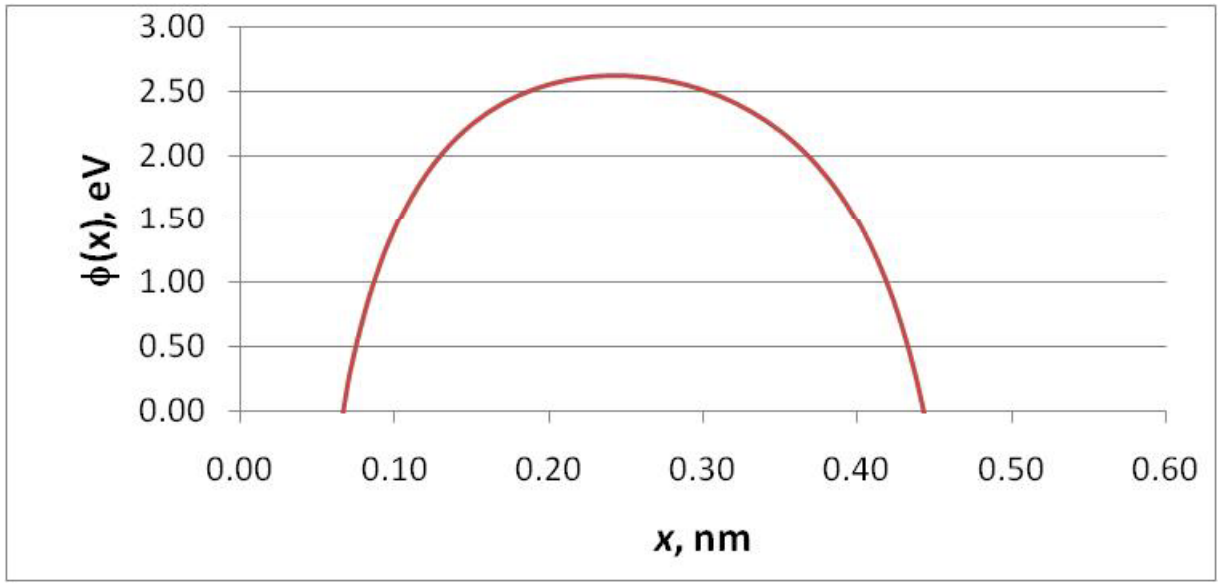


ON/OFF~476

$$a \sim 2n_{Au}^{-\frac{1}{3}} = 0.514nm$$



V=0.5 V



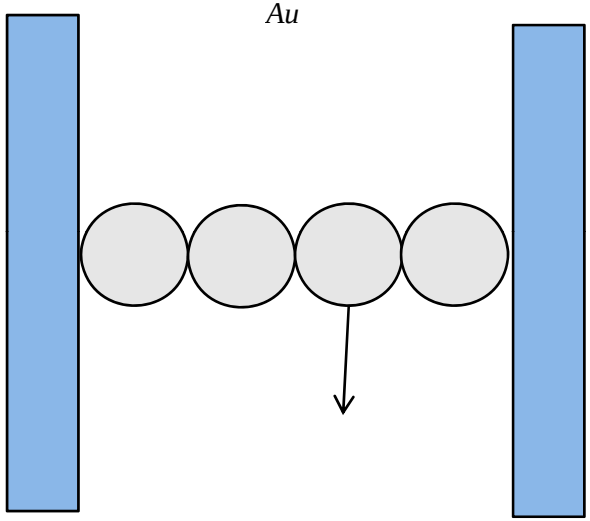
E_b=2.63 eV

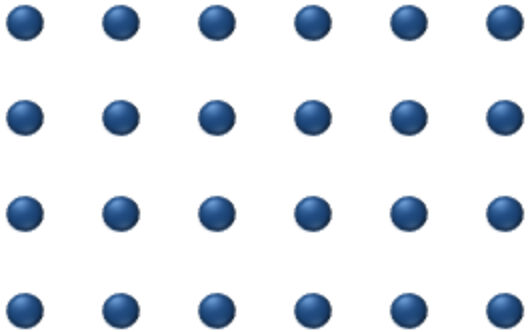
d_t=0.37 nm

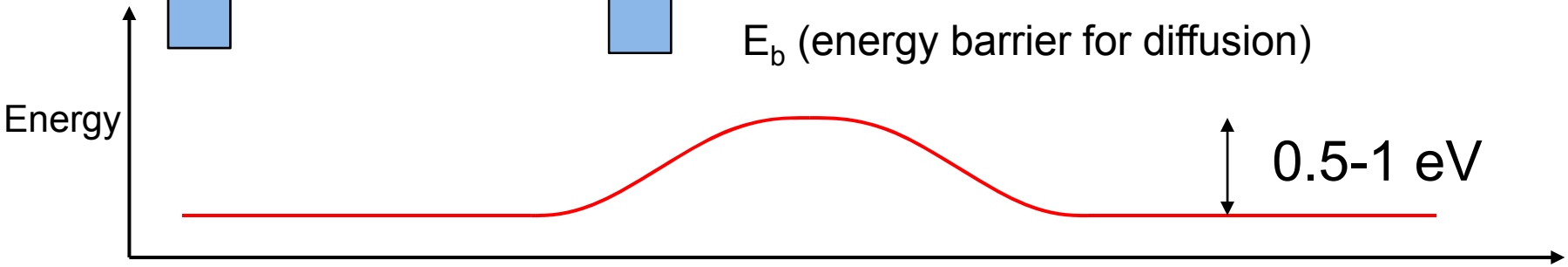
$d_t < a$

Ultimate Atomic Relay: 4-atom gap

$$a \sim 3n_{Au}^{-\frac{1}{3}} \sim 1nm$$



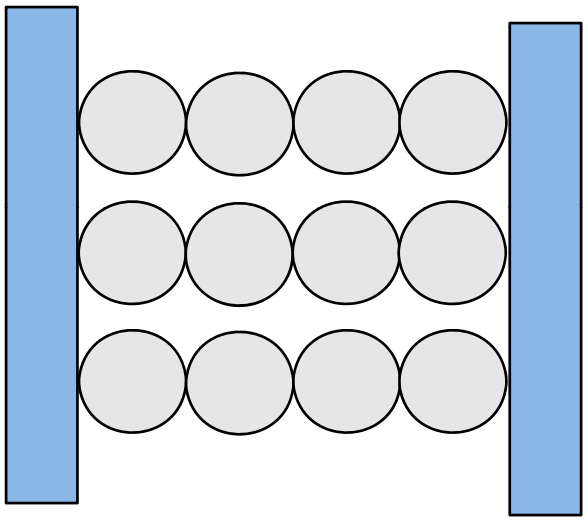


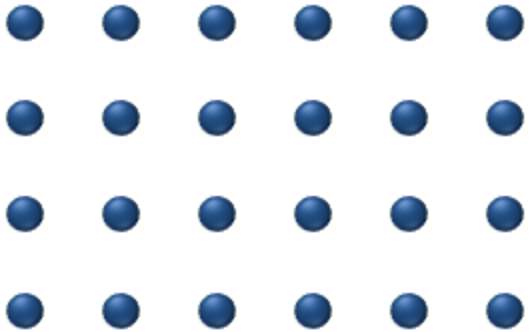


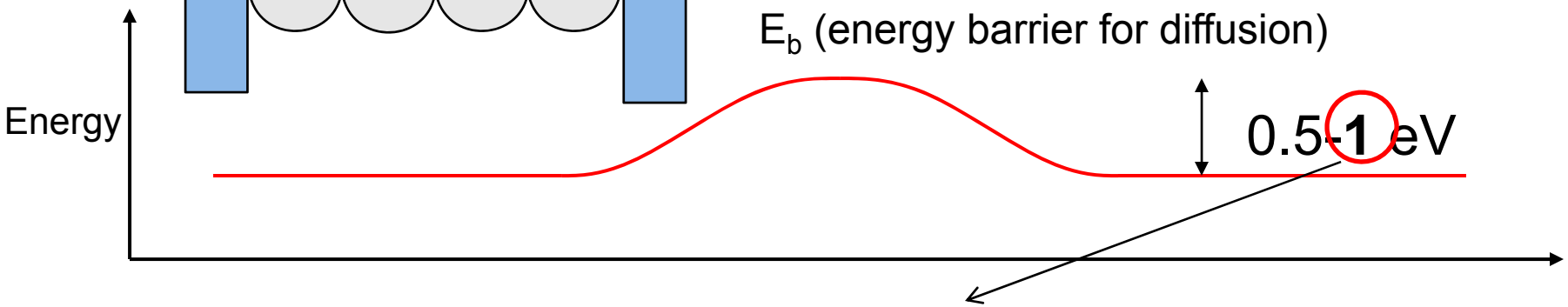
$$t_{tr} = t_0 \exp\left(\frac{E_b}{k_B T}\right) = l_0 \sqrt{\frac{m}{k_B T}} \exp\left[-\frac{E_b}{k_B T}\right] \sim 2s$$

Ultimate Atomic Relay: 4-atom gap

$$a \sim 4n_{Au}^{-\frac{1}{3}} \sim 1nm$$





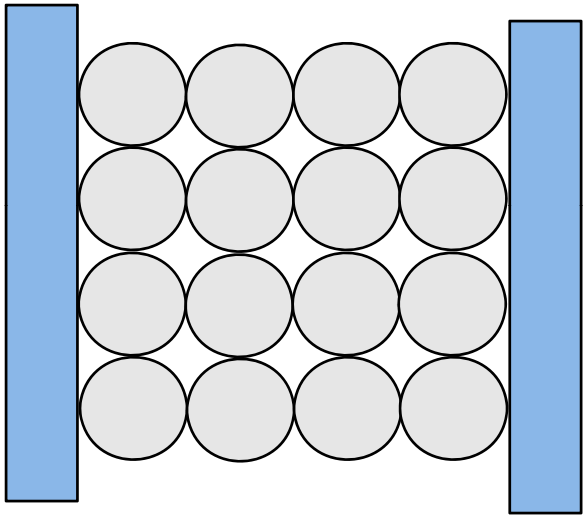


$$t_{tr} = t_0 \exp\left(\frac{E_b}{k_B T}\right)^3 = l_0 \sqrt{\frac{m}{k_B T}} \exp\left[-\frac{E_b}{k_B T}\right]^3 > 10y$$

(E_b=0.5 eV, n>5)

Ultimate ReRAM: A summary

$a = 1nm$



V=1nm^3

N_{at}~100 (64)

E~N_{at}*1eV~10^{-17}J

t~1 ns

E×t×V~10^{-17} J-ns-nm^3

Summary

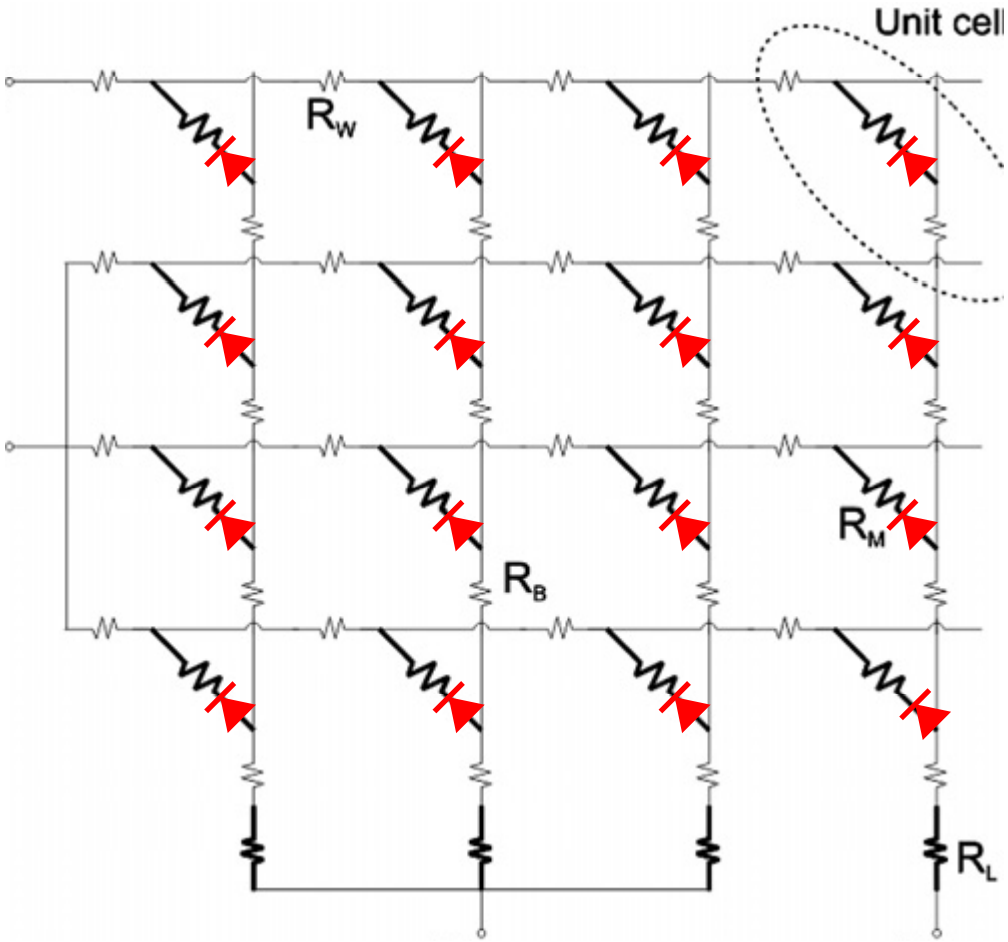
Main constraints due to sensor

	$N_{carriers}$	V, nm^3	E_w, J	t_w, ns	*Space-Action*, J-ns-nm^3	**Biggest component**	
DRAM	10^5	10^5	10^{-14}	1 ns	~10^{-9}	Storage Node	
Flash	10	10^3	10^{-15}	10^3 ns	~10^{-9}	Sensor	FET
STT-RAM	10^5	10^3	10^{-14}	1 ns	~10^{-11}	Selector	FET
ReRAM	100	1	10^{-17}	1 ns	~10^{-17}	Selector	FET

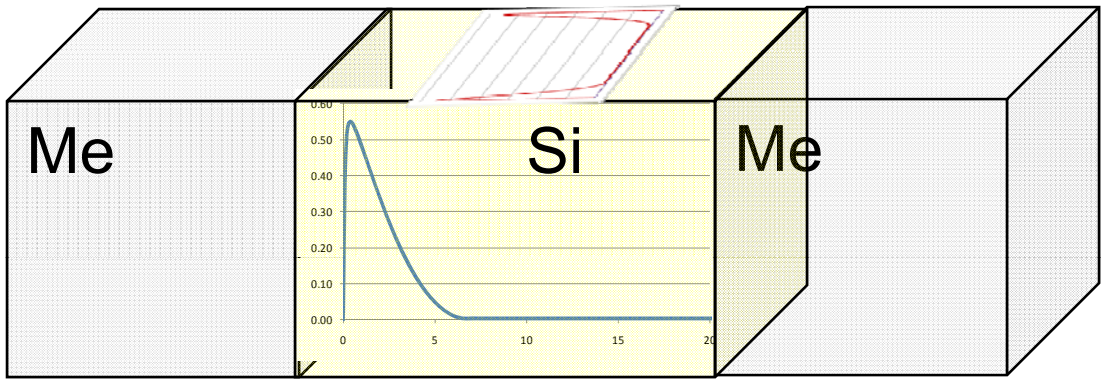
Constraints by sensor not considered

Back-Up slides

FET or Diode selector is biggest part of ReRAM in the limits of scaling

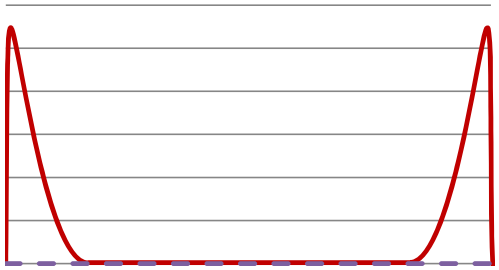


Scaling Limits of Diodes



$N_d \leq N_C$

$$W \approx \sqrt{\frac{2\varepsilon\varepsilon_0 V_{bi}}{N_D}} \sim 10\,nm$$



$N_d \geq N_C$

pn-diode → Esaki tunnel diode

Schottky diode → Ohmic contact

N_C – effective density of states in the conduction band, for Si N_C=2.8x10^{19} cm^{-3}

Space-Action: Flash