The Search for New Information Processing Technologies

Ralph Cavin and Victor Zhirnov

Semiconductor Research Corporation

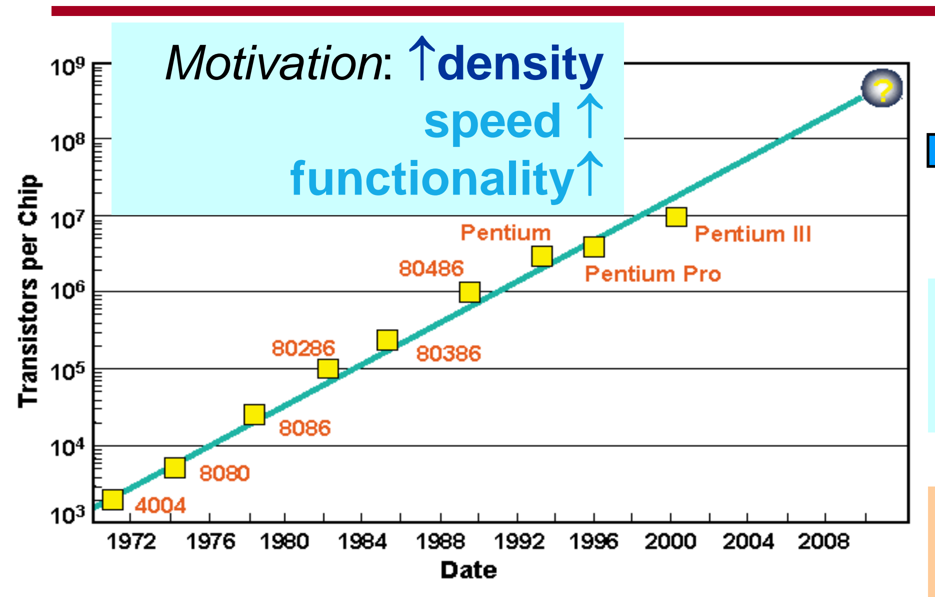
American Physical Society March Meeting 2005

March 22, 2005

International Technology Roadmap for Semiconductors (ITRS)

- A very detailed industrial perspective on the future requirements for nanoscale electronic technologies
  - Goal is to continue exponential gains in performance/price for the next fifteen years
  - Built on worldwide consensus of leading industrial, government, and academic technologists
- Provides guidance for the semiconductor industry and for academic research worldwide
- Content is Technology requirements and Potential Solutions
- Projects that by 2018, half-pitch spacing of metal lines for MPU will be 18 nanometers and MOSFET device gate lengths will be 7 nanometers

Moore's Law: Transistors per chip



Source: Stan Williams, Hewlett Packard

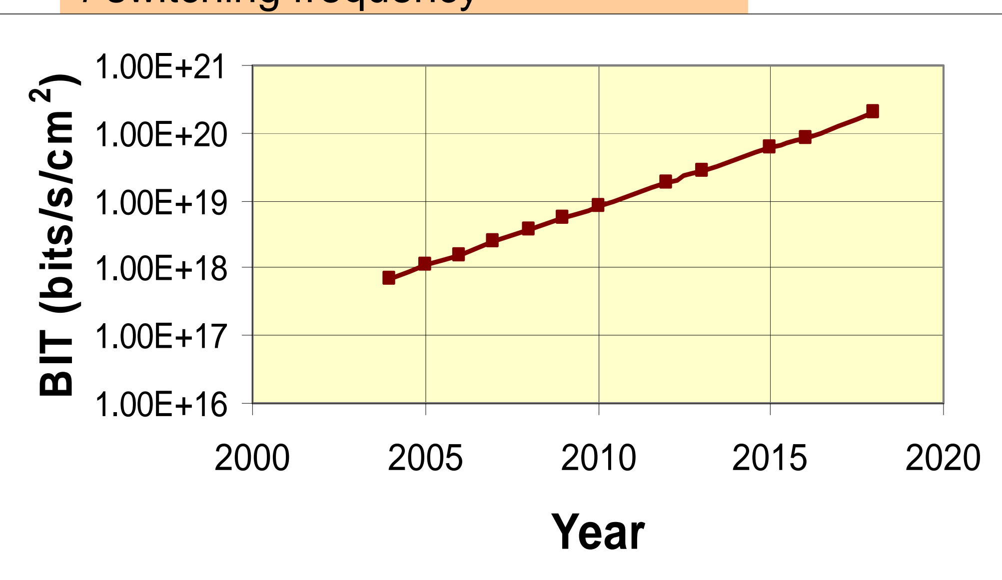
What is the ultimate number of binary transitions per second in a 1cm^2 chip area?

$$BIT = n_{bit} f$$

- a measure of computational capability on device level

n_{bit} – the number of binary states

f-switching frequency



Why scaling? – To increase the *Binary Information Throughput* (BIT)

ITRS CMOS Scaling Challenges –
High-performance Logic Technology Requirements

Year of Production	2010	2012	2013	2015	2016	2018	
Technology Node	*hp45*		*hp32*		*hp22*		
MPU Physical Gate Length (nm)	*18*	*14*	*13*	*10*	*9*	*7*	
Nominal gate leakage current density limit (at 25°C) (A/cm^2)	**1.9E+03**	**2.4E+03**	**7.7E+03**	**1.0E+04**	**1.9E+04**	**2.4E+04**	I_{OFF}
Nominal power supply voltage (V_{dd}) (V)	**1.0**	**0.9**	**0.9**	**0.8**	**0.8**	**0.7**	
Nominal high-performance NMOS sub-threshold leakage current, $I_{sd,leak}$ (at 25°C) (mA/µm)	**0.1**	**0.1**	**0.3**	**0.3**	**0.5**	**0.5**	I_{OFF}
Nominal high-performance NMOS drive current, $I_{d,sat}$ (at V_{dd}, at 25°C) (mA/µm)	**1900**	**1790**	**2050**	**2110**	**2400**	**2190**	I_{ON}
High-performance NMOS intrinsic delay, $\tau = C_{gate} * V_{dd} / I_{d,sat}$ (ps)	**0.39**	**0.30**	**0.26**	**0.18**	**0.15**	**0.11**	I_{ON}
NMOSFET static power dissipation due to drain and gate leakage (W/µm)	**1.10E-06**	**9.90E-07**	**2.97E-06**	**2.64E-06**	**4.40E-06**	**3.85E-06**	I_{OFF}

Energy Costs of Computation: Energy Consumed and Heat generated

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Power Density (Watts/cm²)						
High-performance	155	171	178	—	205	
Power per chip (Watts)						
High-performance	480	530	552		636	

We don't know how to remove that much heat!!

Since each binary transition requires energy E_{bit}, the total power dissipation growth is in proportional to the information throughput:

$$P = \frac{n_{bit}}{t_{sw}} \cdot E_{bit} = BIT \cdot E_{bit}$$

A universal relation for information processing devices

Two trivial solutions:

(1) reduce E_{bit} (operation voltage)
 - $V_{min} > k_B T \ln 2 = 17$ mV *(signal-to-noise ratio/distinguishability)*

(2) reduce number of binary transitions, e.g. by decreasing the activity factor of the transistors in an IC
 - equivalent to stopping scaling

Technology Enhancements

High Performance Ultra-Scaled Silicon CMOS

I_{on}/I_{off} is regarded by many as the primary driver for the selection of the future candidates for logic devices (binary switches)

I_{on}=speed

I_{off}=power&error

This applies to all charge-based devices

Calculations performed using MASTAR – ST Microelectronics – T. Skotnicki

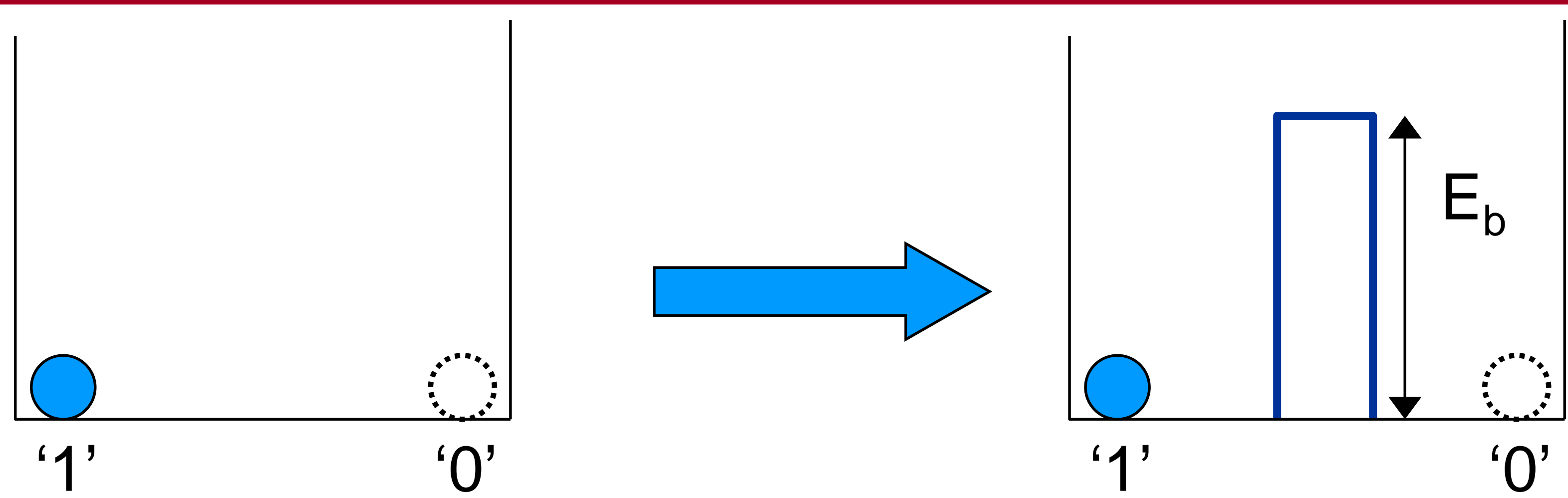
Emerging Research Logic Devices

2003 ITRS PIDS/ERD Chapter

Device	*FET*	*RSFQ*	*1D structures*	*Resonant Tunneling Devices*	*SET*	*Molecular*	*QCA*	*Spin transistor*
Cell Size	100 nm	0.3 μm	100 nm	100 nm	40 nm	Not known	60 nm	100 nm
Density (cm^{-2})	3E9	1E6	3E9	3E9	6E10	1E12	3E10	3E9
Switch Speed	700 GHz	1.2 THz	Not known	1 THz	1 GHz	Not known	30 MHz	700 GHz
Circuit Speed	30 GHz	250–800 GHz	30 GHz	30 GHz	1 GHz	<1 MHz	1 MHz	30 GHz
Switching Energy, J	2×10^{-18}	$>1.4\times10^{-17}$	2×10^{-18}	$>2\times10^{-18}$	$>1.5\times10^{-17}$	1.3×10^{-16}	$>1\times10^{-18}$	2×10^{-18}
Binary Throughput, $GBit/ns/cm^2$	86	0.4	86	86	10	N/A	0.06	86

We HAVE IDENTIFIED NO VIABLE EMERGING LOGIC TECHNOLOGIES for Information Processing beyond CMOS

Energy Barriers in Electronic Devices (FAST SWITCH PROBLEM)

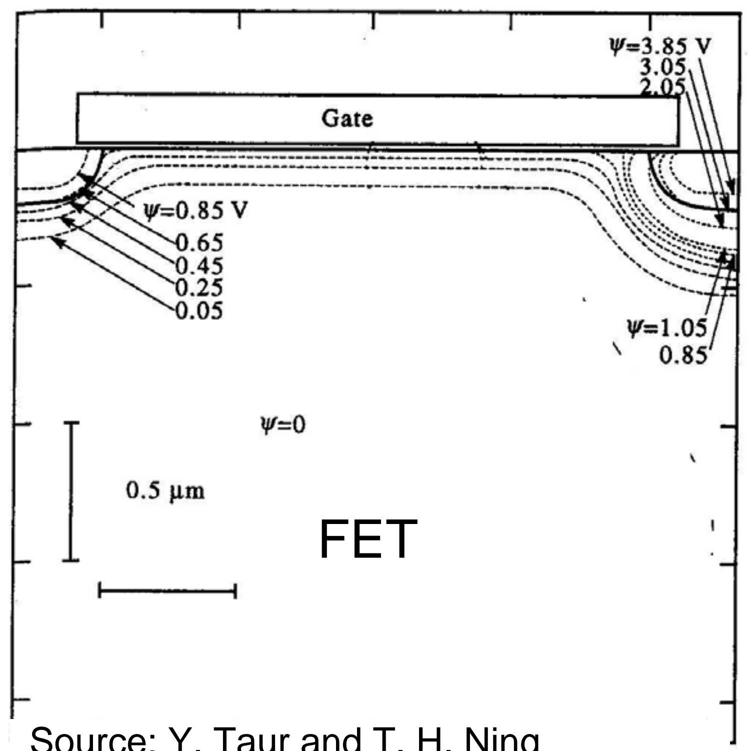


This structure cannot be used for representation/processing information

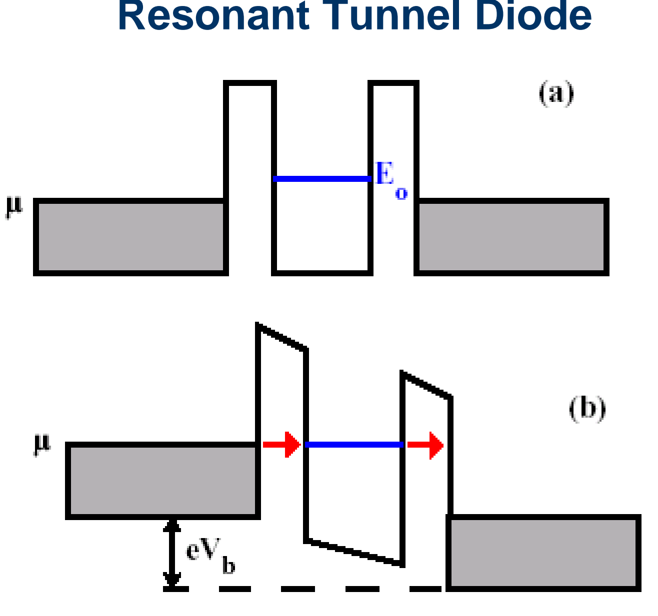
An energy barrier is needed to preserve a binary state

Energy Barriers in Materials

- Any electronic device contains at least one energy barrier, which controls electron flow. The barrier properties, such as height, length, and shape determine the characteristics of electronic devices.

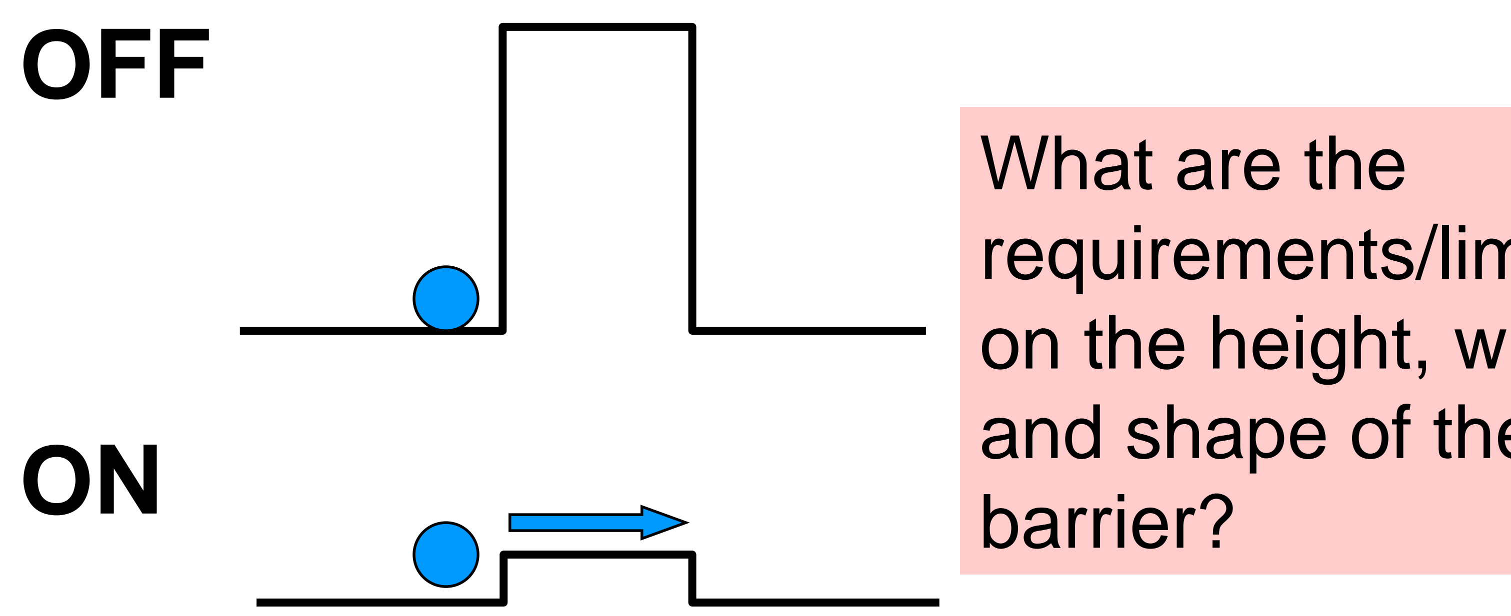


Source: Y. Taur and T. H. Ning
"Fundamentals of modern VLSI devices"
Cambridge university Press 1998



R. Compano (Ed.)
Technology Roadmap for Nanoelectronics
(European Communities, 2001)

Arbitrary Electronic Device



What are the requirements/limitations on the height, width and shape of the barrier?

Classic Distinguishability: *The Boltzman constraint*

How small the energy barrier height could be ?

Barrier control (gate)

a

E_b

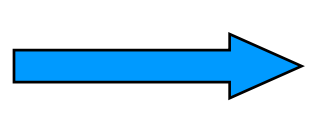
a

E_b

OR

E_b

$$E_{bit}^{min} = E_b$$



Distinguishability requirement: The probability of spontaneous transitions(errors) Π_{error}<0.5 (50%)

$$\Pi_{error} = \exp(\frac{E_b}{k_B T})$$

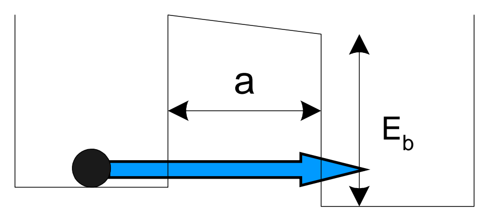
$$0.5 = \exp(-\frac{E_b}{k_B T})$$

$$E_b^{min} = k_B T \ln 2$$

Quantum Distinguishability: The Heisenberg Constraint

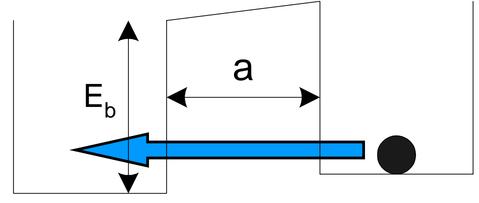
$$\Delta x \Delta p \geq \hbar$$

$$a_{crit} = \frac{\hbar}{\sqrt{2mE_b}}$$



$$E_b^{\min} \sim \frac{\hbar^2}{2ma^2}$$

(@Π=0.14)



Wigner-Kramers-Brillouin (WKB) approximation for tunneling:

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

$$E_b^{\min} = \frac{\hbar^2 \ln^2 2}{8ma^2}$$

(@Π=0.5)

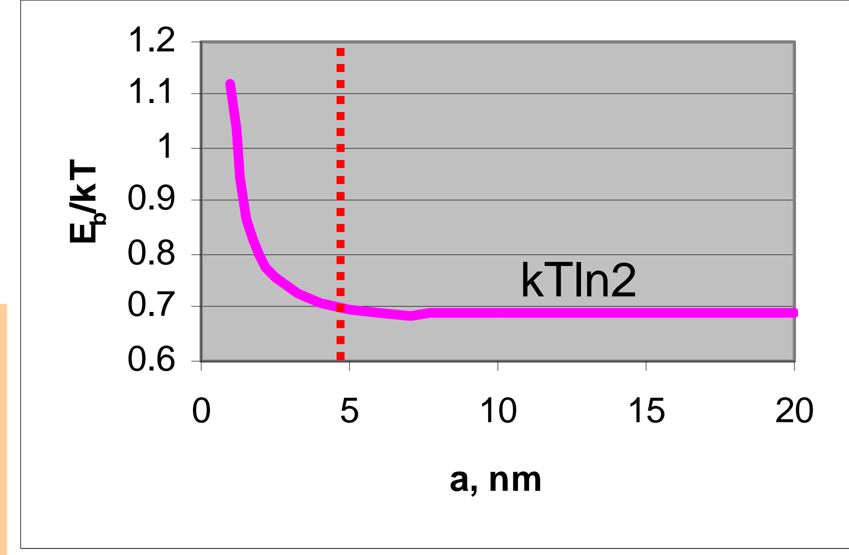
Total Distinguishability @ Π=0.5

$$\Pi_{error} = \Pi_{classic} + \Pi_{quantum} - \Pi_{classic}\Pi_{quantum} =$$

$$= \exp(-\frac{E_b}{kT}) + \exp(-\frac{2\sqrt{2m}}{\hbar} a\sqrt{E_b}) - \exp(-\frac{\hbar E_b + 2akT\sqrt{2mE_b}}{\hbar kT})$$

Generalized expression for the minimum energy barrier to create a bit

$$E_b^{\min} \approx kT\ln 2 + \frac{\hbar^2(\ln 2)^2}{8ma^2}$$



What does fundamental physics have to say about electron transport limits?

1) Minimum energy per binary transition

$$E_{bit}^{\min} = k_B T \ln 2$$

2) Minimum distance between two distinguishable states (Heisenberg)

$$\Delta x \Delta p \geq \hbar \;\rightarrow\; x_{\min} = a = \frac{\hbar}{\sqrt{2mkT\ln 2}} = 1.5nm(300K)$$

3) Minimum state switching time (Heisenberg)

$$\Delta E \Delta t \geq \hbar \;\rightarrow\; t_{st} = \frac{\hbar}{kT\ln 2} = 4\times 10^{-14}\, s(300K)$$

4) Maximum gate density

$$n = \frac{1}{x_{\min}^2} = 4.6\times 10^{13}\,\frac{gate}{cm^2}$$

Total Power Dissipation (@E_{bit}= $kT\ln(2)$)

$$P_{chip} = \frac{n \cdot E_{bit}}{t} = 4.6 \cdot 10^{13}[cm^{-2}] \cdot \frac{3 \cdot 10^{-21}[J]}{4 \cdot 10^{-14}[s]}$$

$$E_{bit} = k_B T \ln 2 \approx 3 \cdot 10^{-21} J$$

$$P_{chip} = 4.74 \times 10^6 \frac{W}{cm^2}$$

T=300 K

The circuit would vaporize when it is turned on!

How much heat a solid system can tolerate?…

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Power Density (Watts/cm²)						
High-performance	**155**	**171**	**178**	**—**	**205**	
Power per chip (Watts)						
High-performance	**480**	**530**	**552**		**636**	

We don't know how to remove that much heat!!

***Several hundred* W/cm²** is close to known limits of heat removal from a 2-dimensional solid material structure with **T_{max} =125°C**

Experimental demonstrations of on-Si cooling systems (without active devices):

680 W/cm² thermoelectric (Zheng et al.)

790 W/cm² microchannel (Tuckerman and Pease)

What are the fundamental limits for heat removal?

Should we consider cryogenic devices as an option for the mainstream applications?

Three fundamentals of heat removal:

1)The Newton's Law of Cooling: $q=h(T_h-T_a)$

(h-heat transfer coefficient)

2) The Ambient: **T_a=300 K !!!**

3) The Carnot's theorem:

$$W_{cool} = \frac{T_a - T_c}{T_c} Q^-$$

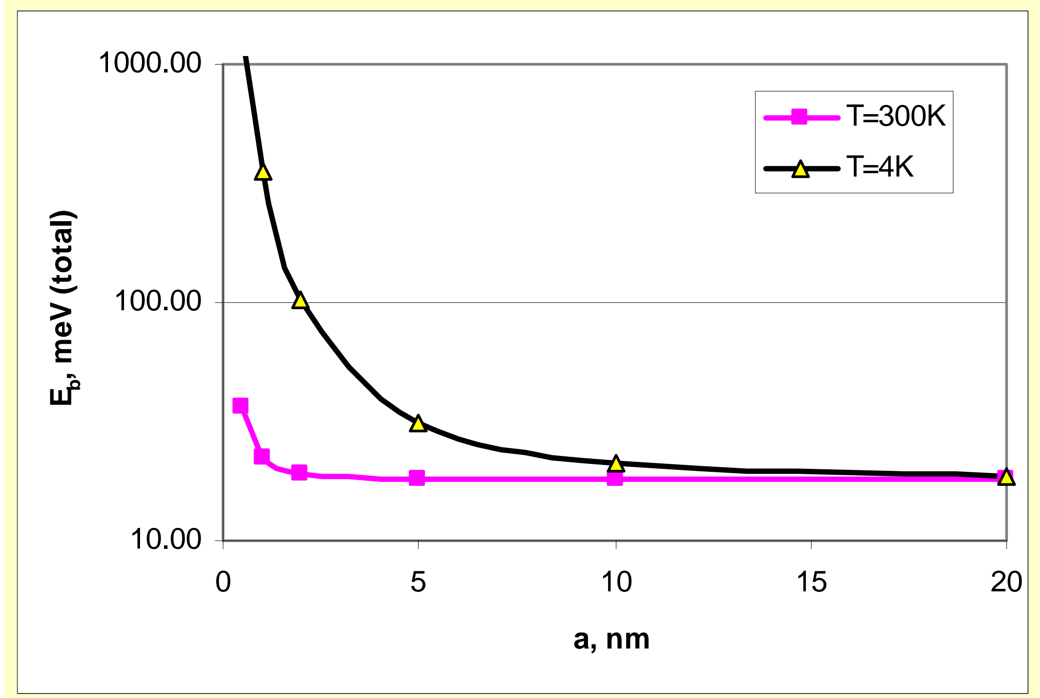
Work to be done

Heat to be removed

Cryogenic Computation with Nanodevices

$$E_{bit}^{total} = E_{bit} + \frac{T_a - T_{dev}}{T_{dev}} E_{bit} = \frac{T_a}{T_{dev}} E_{bit} = \frac{T_a}{T_{dev}} \left[k_B T_{dev} \ln 2 + \frac{\hbar^2 (\ln 2)^2}{8ma^2} \right]$$

$$E_{bit}^{total} = k_B T_a \ln 2 + \frac{T_a}{T_{dev}} \frac{\hbar^2 (\ln 2)^2}{8ma^2} > k_B T_a \ln 2$$



Due to tunneling, the power consumed by the device depends on both operating temperature and size that manifests itself with unexpectedly dramatic increases in total power consumption at cryogenic temperatures.

Asymptotically Dissipation-less Computing?

Often referred as to "reversible" or "adiabatic" computing…

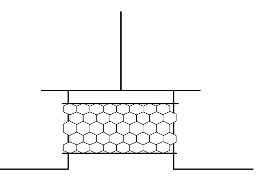
We aren't optimistic about Reversible or Adiabatic Computing:

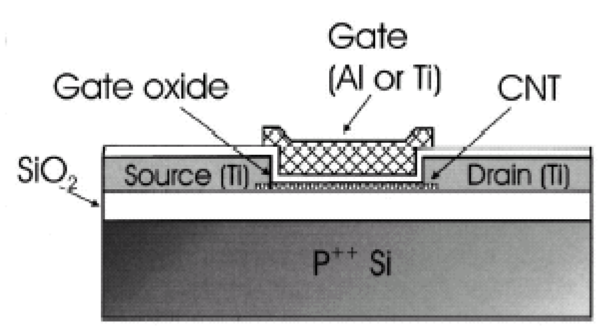
- Reversible Computing requires isolation from external environment
  - Cooling to very low temperature is very costly in terms of energy
- "Adiabatic" methods attempt to save and re-use
  - All proposed methods are inefficient
  - Require slow device operation for which there is limited interest

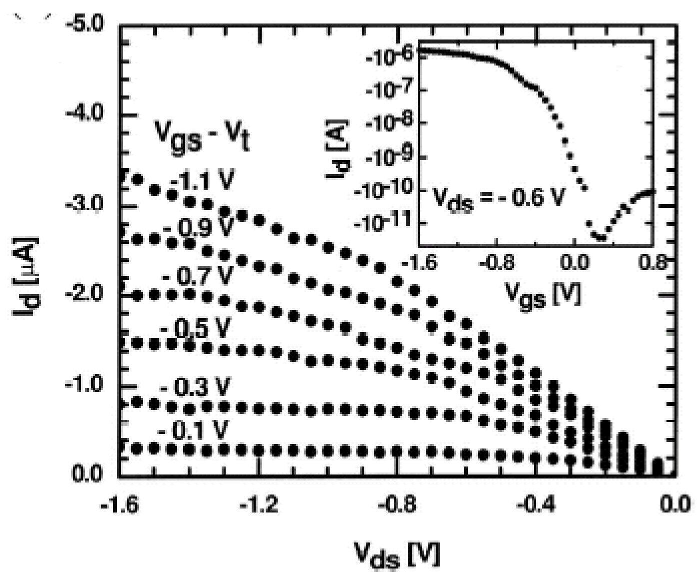
Some Current Ideas for Future Logic Devices

CNT and Nanowire transistor
Resonant tunneling devices
Single electron transistor
Molecular transistor
Spintronics

CNT transistor







S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and Ph. Avouris, Appl. Phys. Lett 80 (2002) 3817

Questions for CNT FETs

1) Can CNT FET be smaller, faster and dissipate less energy than Si FET?

2) Is it possibilities to integrate individual CNT components in a complex circuit (billions of components per cm^2)?

3) Is Ballistic Transport a big advantage?

$$v_{drift} = \mu \cdot F$$

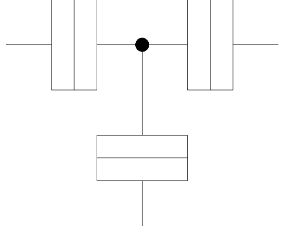
$$j = e \cdot n \cdot v$$

$n_{drift} > n_{bal}$

$$v_b = \left(\frac{e}{m} t\right) \cdot F$$

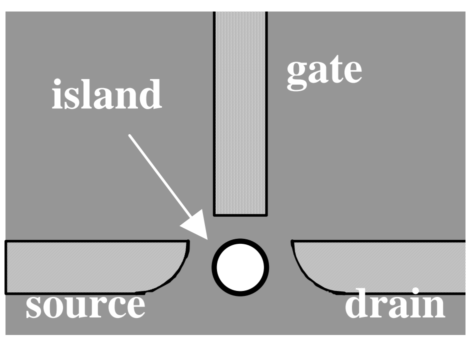
For long channels, is ballistic transport possible in the high-current regime?

Single electron transistor



Single electron transistor (SET)

- **Electron movements are controlled with single electron precision**
- **Tunneling and Coulomb blockade**



Single-Electron Transistor has all problems of charge-based devices

The Fan-Out requirement is not satisfied?

High error rate?

Low speed?

What about FET in single-electron mode? (FET will be 32-electron transistor by 2018)

Resonant tunneling devices(diodes and transistors)

This year (2004) marks the 30 years of experimental studies of resonant tunneling structures (CHANG LL, ESAKI L, TSU R / IBM, “RESONANT TUNNELING IN SEMICONDUCTOR DOUBLE BARRIERS”, APPL. PHYS. LETT 24 (12): 593-595 1974)

1974	Chang
1986	Reed
1986	Reed
1988	Broekaert
1990	Broekaert
1990	Mehdi
1991	Chen
1991	Smet
1992	Watanabe
1995	Moise
1999	Miyamoto
2000	Watanabe
2001	Ishikawa
2002	Kikuchi
2002	Kado
2002	Wernersson
2002	Malindretos
2002	Bjork
2003	Wang
2003	Ikeda
2004	Xu

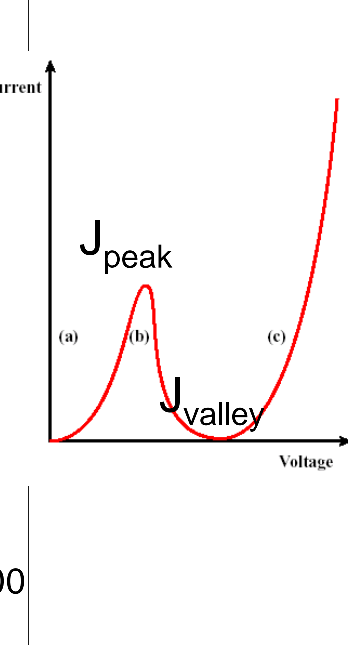
FET

1.00E+07
1.00E+05
1.00E+03
1.00E+01
1.00E-01
1.00E-03
1.00E-05
1.00E-07

Jpeak, A/cm2

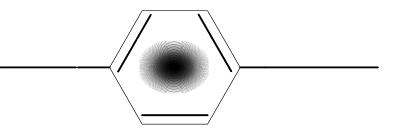
1 10 100 1000 10000 100000 1000000

Peak-to-Valley Ratio



Can RTDs match FETs both in I_{ON} and I_{OFF}?

Molecular transistor



- Molecular Transistor has all problems of charge-based devices
- Question: Can molecular transistor be smaller, faster and dissipate less energy than Si FET?
- Many fundamental issues in integration and fabrication

Spin transistors: Spin Transport vs. Charge Transport?

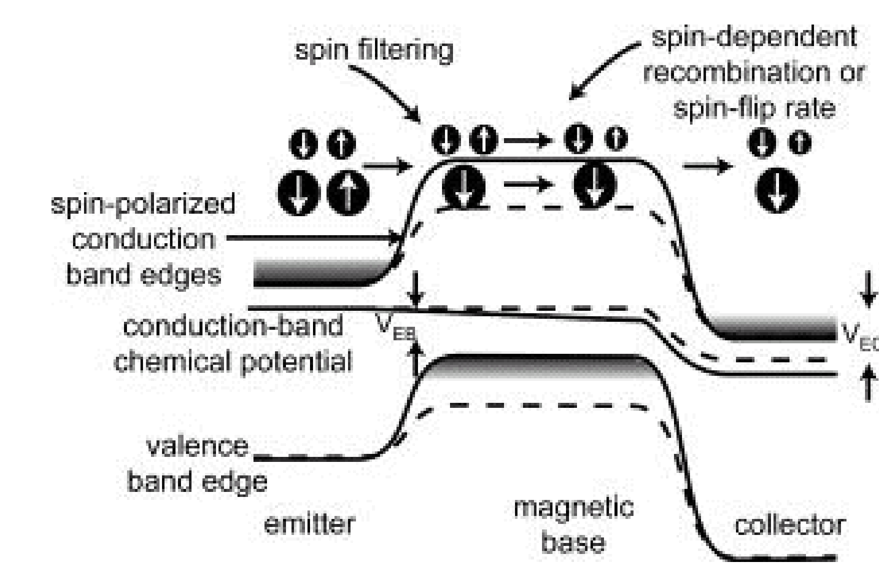
- Spin is a property of material particles (e.g. electron, proton etc.)
- To move spin from point A to point B requires moving material particle
- *Question: Even if we are controlling spin we are still moving electrons. Don't we have the same problems as with charge-based devices?*

Examples of proposed Spin transistors:

- Johnson tranistor
- Miziushima FIFS transistor
- SPICE transistor
- Ounadjela-Hehn transistor
- Datta-Das transistor…

Operation principle: control of spin-polarized flow of electrons

- Spin-valve (GMR)
- Magnetic tunnel junction
- Spin Transport in 2DEG



Flatte ME, Yu ZG, Johnston-Halperin E, et al.
Theory of semiconductor magnetic bipolar transistors,
APPL. PHYS. LETT. 82 (2003) 4740

Which of current nanoelectronic concepts will become the NEW SWITCH?

ANSWER:

NONE YET

What else could be done?
Emerging Memory Devices

Invent high speed, high density, ***electrically accessible, non-volatile*** memory**!**

Memory and Logic in a typical computer system

How can NVM enter this hierarchy?

Memory	Level	Properties
SRAM	1st level cache	Density-low, Capacity-low, Speed-high, Volatile
SRAM	2nd level cache	Density-low, Capacity-medium, Speed-high, Volatile
DRAM	Main memory	Density-high, Capacity-high, Speed-high, Volatile
	Hard disk drive	Density-medium, Capacity-high, Speed-low, Non-Volatile

CPU

Memory limits computer's performance

- Overall computer's ability to execute programs are limited by interaction between MPU and memory
- This problem is not automatically solved by scaling
- Evolutionary Solution: SRAM occupies increasing amount of floorspace of MPU chip – Less room for LOGIC
  - Decrease the net information throughput
- Volatility of semiconductor memory requires external storage media with slow access (magnetic hard drives, optical CD)
- Invention of *electrically accessible* *non-volatile* memory with *high speed* and *high density* would imply a **revolution in computer architectures**

New Memory Hierarchy ?

CPU

1st level cache

2nd level cache

Main memory

Hard disk drive

?

CPU

Density-high
Capacity-high
Speed-high
Non-Volatile

?

CPU

Density-high
Capacity-high
Speed-high
Non-Volatile

Emerging Research Memory Devices

2003 ITRS PIDS/ERD Chapter

Storage Mechanism	*Present Day Baseline Technologies*		*Phase Change Memory**	*Floating Body DRAM*	*Nano-floating Gate Memory***	*Single/Few Electron Memories***	*Insulator Resistance Change Memory***	*Molecular Memories***
Device Types	DRAM	NOR Flash	OUM	1TDRAM eDRAM	Engineered tunnel barrier or nanocrystal	SET	MIM oxides	Bi-stable switch Molecular NEMS
Availability	2004	2004	~2006	~2006	~2006	>2007	~2010	>2010
Cell Elements	1T1C	1T	1T1R	1T	1T	1T	1T1R	1T1R

Several Emerging Memory Technologies Show Promise

What Else Could be Done? – Explore new approaches

- “Different” spintronics

(1) Can we make spin devices that operate without moving electrons?
(2) Non-equilibrium operation of binary switch?

- Orbitronics

Interplay between charge, orbital, and spin degrees of freedom. Materials with both ferroelectricity and ferromagnetism (electrically controlled)

Binary State 0
Positive Electric Field
Positive Polarization

Binary State 1
Negative Electric Field
Negative Polarization

(1) Can we move atoms instead of moving electrons?

- Phononics: thermal breakthrough

(1) How can we better remove heat and what are the fundamental limits of heat removal?
(2) How can we isolate selected materials subsystems from thermal noise?
(3) Are there ways to control phonon movement by external stimuli.
Co-design of electric and thermal circuits?

- Invent high speed, high density, *electrically accessible, non-volatile* **memory**.