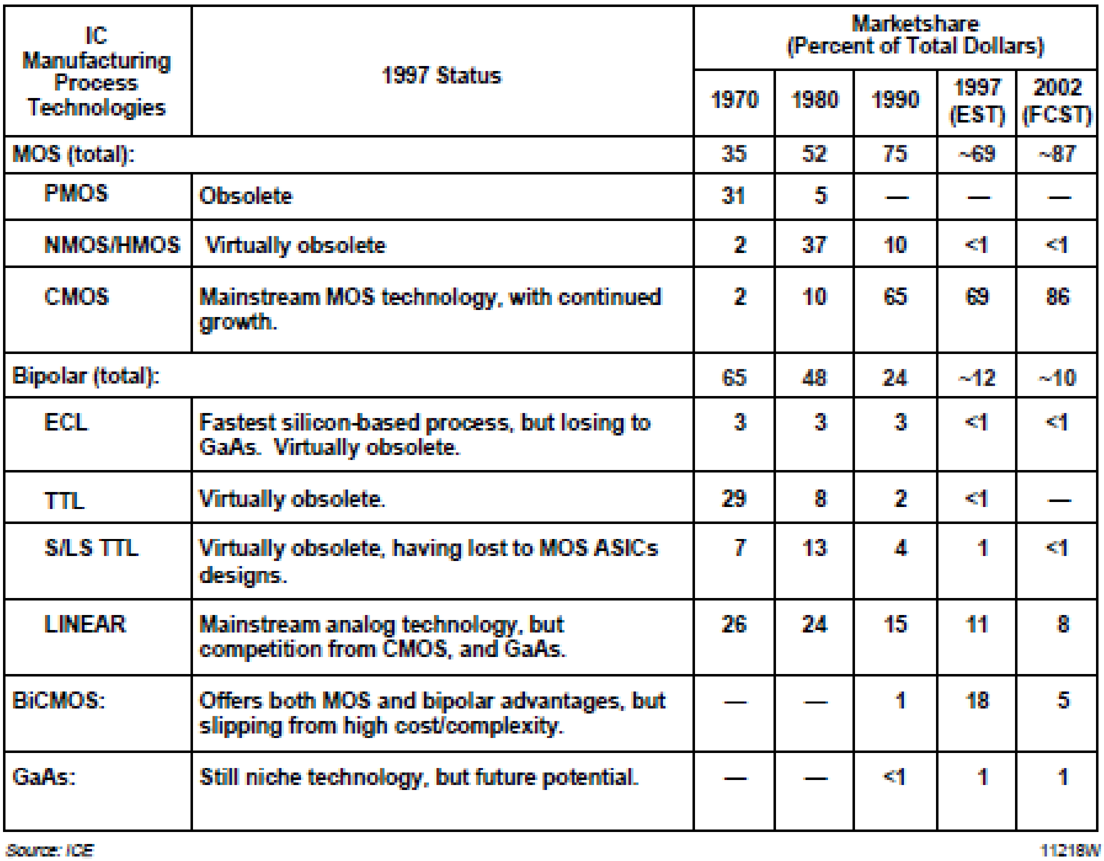
Trends in Integrated Circuits Technology

IC DEVICE TECHNOLOGY OVERVIEW

There are a variety of major manufacturing process technologies (Figure 4-1) used in design and fabrication of silicon-based integrated circuits (ICs). These include metal-oxide-semiconductor (MOS), bipolar, and combined bipolar and complementary-MOS (BiCMOS). While silicon-based processing dominates in semiconductor manufacturing, gallium arsenide (GaAs), a compound-semiconductor material, is a niche alternative to silicon for some applications.



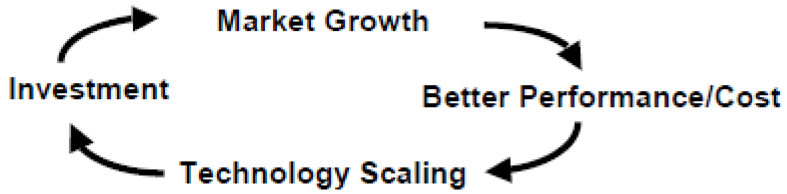
IC Manufacturing Process Technologies	1997 Status	Marketshare (Percent of Total Dollars)				
		1970	1980	1990	1997 (EST)	2002 (FCST)
MOS (total):		35	52	75	~69	~87
PMOS	Obsolete	31	5	—	—	—
NMOS/HMOS	Virtually obsolete	2	37	10	<1	<1
CMOS	Mainstream MOS technology, with continued growth.	2	10	65	69	86
Bipolar (total):		65	48	24	~12	~10
ECL	Fastest silicon-based process, but losing to GaAs. Virtually obsolete.	3	3	3	<1	<1
TTL	Virtually obsolete.	29	8	2	<1	—
S/LS TTL	Virtually obsolete, having lost to MOS ASICs designs.	7	13	4	1	<1
LINEAR	Mainstream analog technology, but competition from CMOS, and GaAs.	26	24	15	11	8
BiCMOS:	Offers both MOS and bipolar advantages, but slipping from high cost/complexity.	—	—	1	18	5
GaAs:	Still niche technology, but future potential.	—	—	<1	1	1

Source: ICE

11218W

Figure 4-1. Market Share Overview of IC Manufacturing Process Technologies

Miniaturization => Market growth

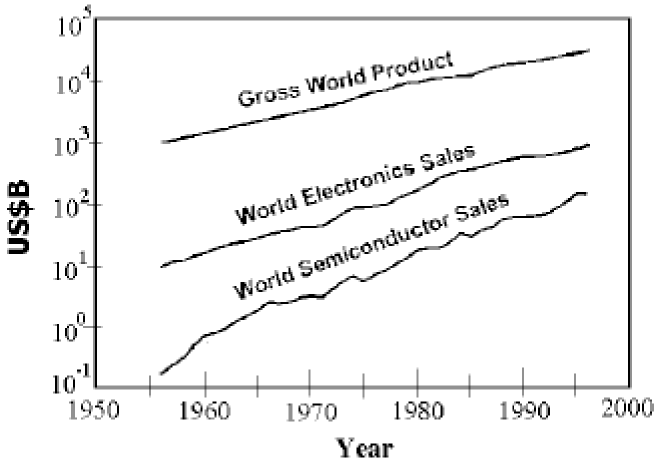


Market Growth

Better Performance/Cost

Technology Scaling

Investment

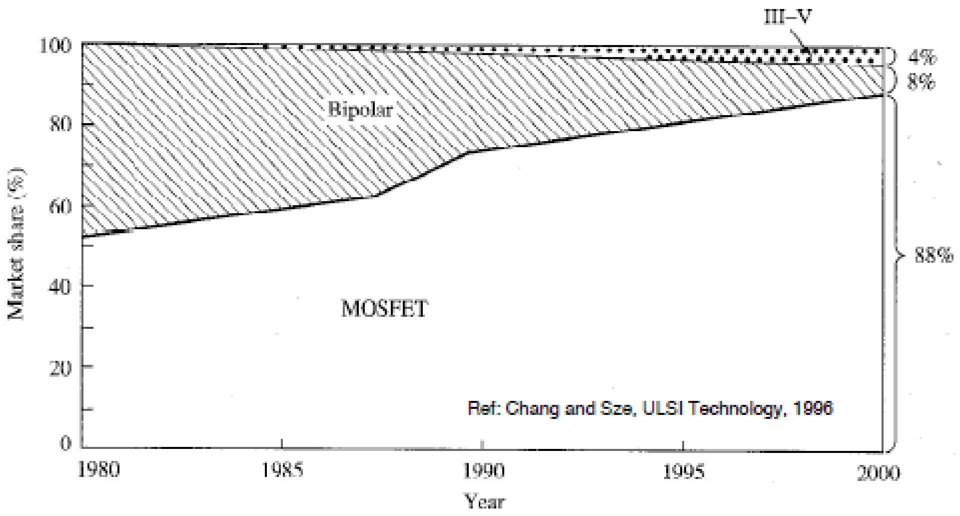


Semiconductors have become increasingly more important part of world economy

Courtesy Prof. Tsu-Jae King

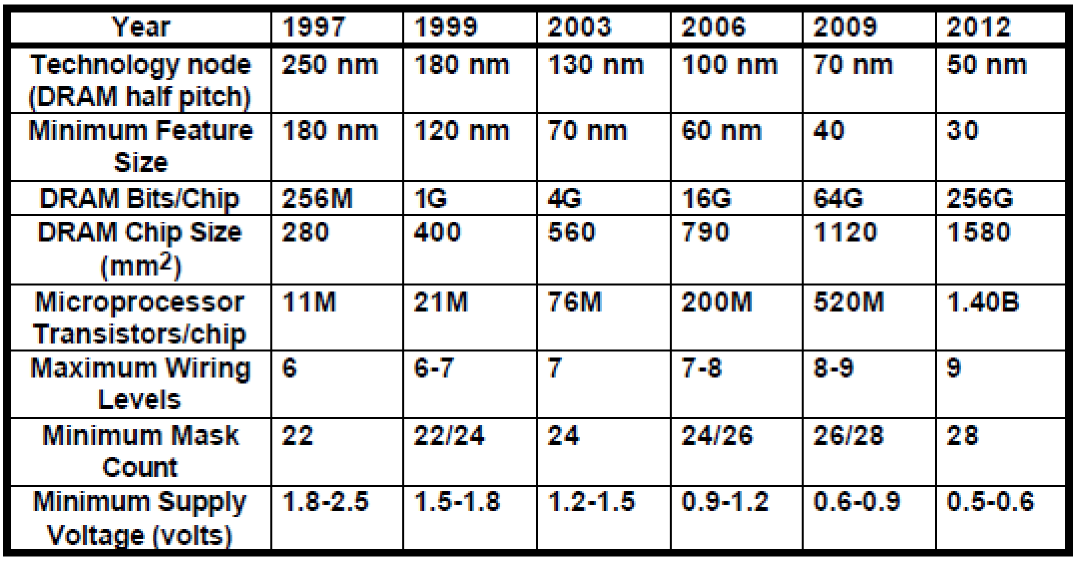
(Sources: VLSI Research Inc.; United Nation yearbook; World Bank Database; IMF)

World IC Market by Technology



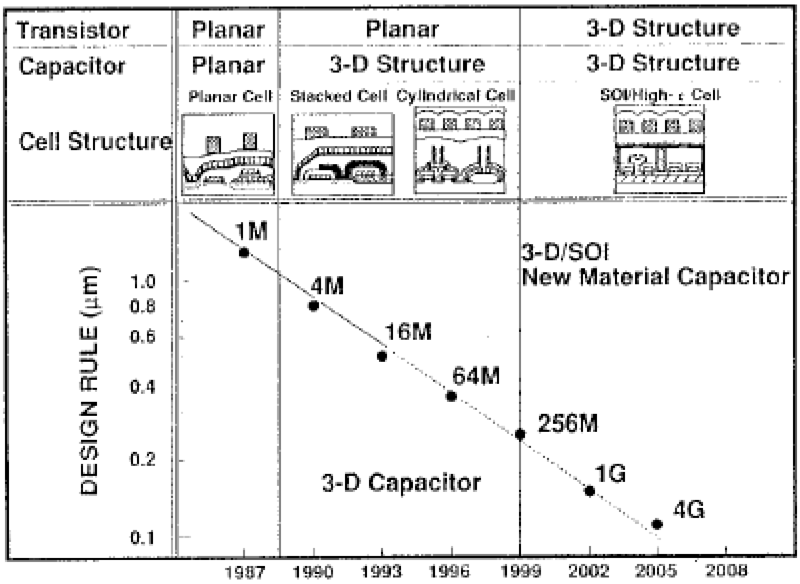
Ref: Chang and Sze, ULSI Technology, 1996

Silicon CMOS has become the pervasive technology



Year	1997	1999	2003	2006	2009	2012
Technology node (DRAM half pitch)	250 nm	180 nm	130 nm	100 nm	70 nm	50 nm
Minimum Feature Size	180 nm	120 nm	70 nm	60 nm	40	30
DRAM Bits/Chip	256M	1G	4G	16G	64G	256G
DRAM Chip Size (mm^2)	280	400	560	790	1120	1580
Microprocessor Transistors/chip	11M	21M	76M	200M	520M	1.40B
Maximum Wiring Levels	6	6-7	7	7-8	8-9	9
Minimum Mask Count	22	22/24	24	24/26	26/28	28
Minimum Supply Voltage (volts)	1.8-2.5	1.5-1.8	1.2-1.5	0.9-1.2	0.6-0.9	0.5-0.6

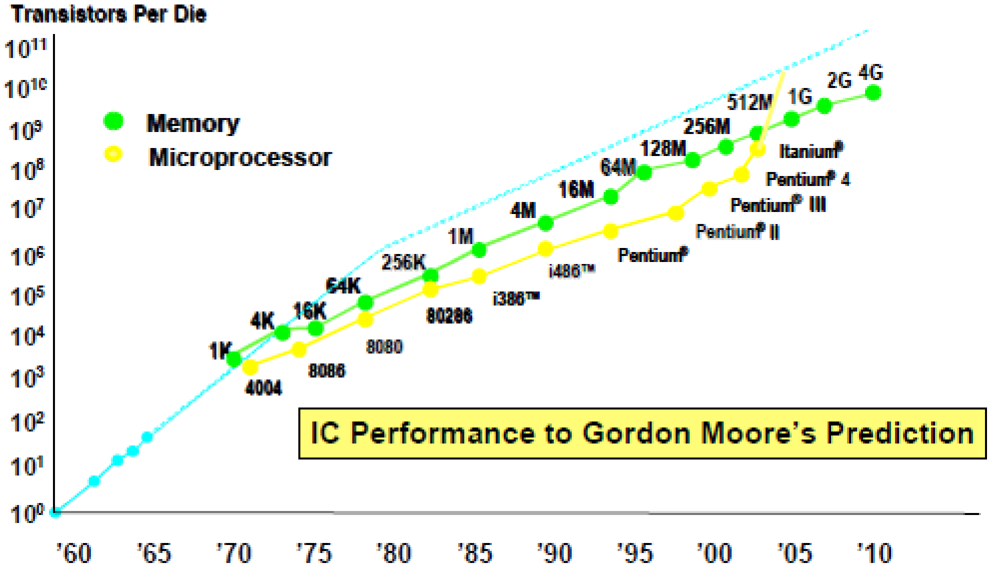
Future projections for silicon technology taken from the SIA ITRS 1999



Ref. H. komiya IEEE ISSCC 1993

Device structures are becoming increasingly more complex

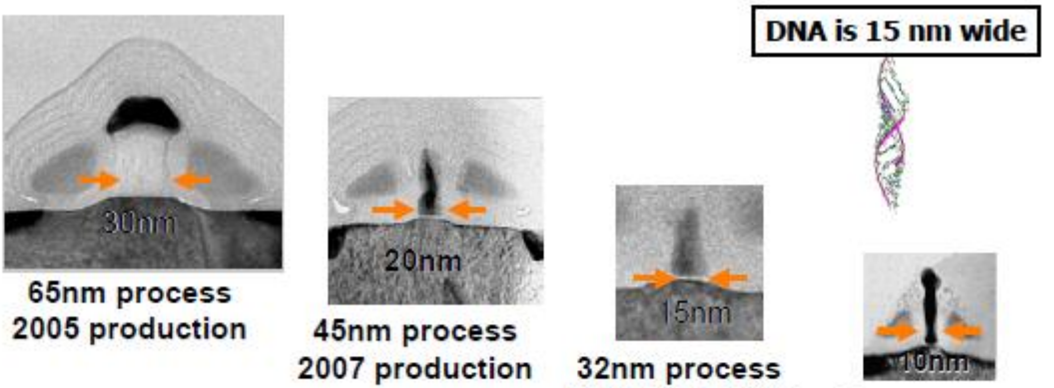
Moore's Law



Source: Intel

Intel's Transistor Research down to 10nm

Electronics is Nanotechnology



65nm process
2005 production

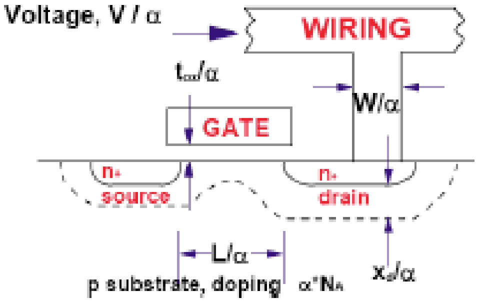
45nm process
2007 production

32nm process
2009 production

22nm process
2011 production

Source: Intel

MOS Device Scaling



Constant E Field Scaling

All device parameters are scaled by the same factor α.

- Gate oxide thickness t_{ox} ↓
- Channel length L ↓
- Source/drain junction depth X_j ↓
- Channel doping ↑
- Supply voltage V_D ↓

Why do we scale MOS transistors?

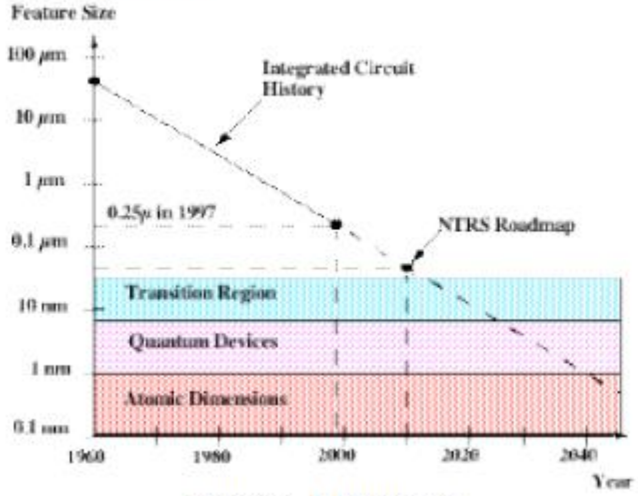
1. Increase device packing density ~ α^2
2. Improve frequency response (speed) ~ a
3. Power/ckt: ~$1/\alpha^2$, power density constant
4. Improve current drive (transconductance g_m)

$$g_m = \left.\frac{\partial I_D}{\partial V_G}\right|_{V_D = const}$$

$$= \frac{W}{L}\mu_n\frac{K_{ox}}{t_{ox}}V_D \qquad \text{for } V_D < V_{D_{SAT}},\ \textit{linear region}$$

$$= \frac{W}{L}\mu_n\frac{K_{ox}}{t_{ox}}(V_G - V_T) \quad \text{for } V_D > V_{D_{SAT}},\ \textit{saturation region}$$

How far can we continue to scale?

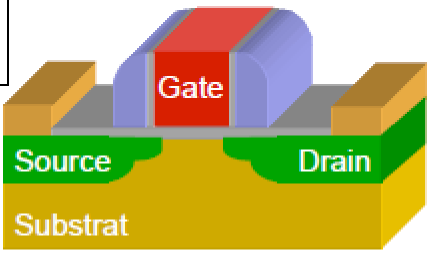


(Source: J. Plummer)

Physical Limits in Scaling Si MOSFET

Gate stack
- Tunneling current ⇒ Increased I_{off}
- Gate depletion ⇒ Increased EOT

Source/Drain
- Contact resistance
- Doping level, abruptness

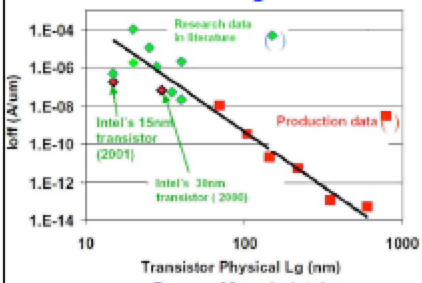


High E-Field
- Mobility degradation
- Reliability

Channel
- Surface scattering - the “universal mobility” tyranny
- DIBL ⇒ drain to source leakage
- Subthreshold slope limited to 60mV/decade (kT/q) ⇒ Increased I_{off}
- V_G - V_T decrease ⇒ reduced I_{ON}

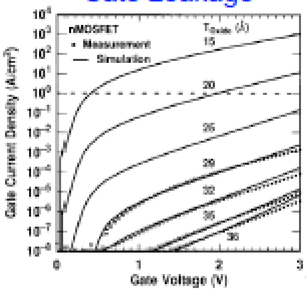
MOSFET Scaling Limit: Leakage

Total Leakage Trend



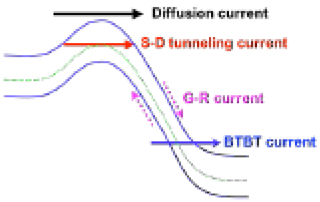
Source: Marcyk, Intel

Gate Leakage

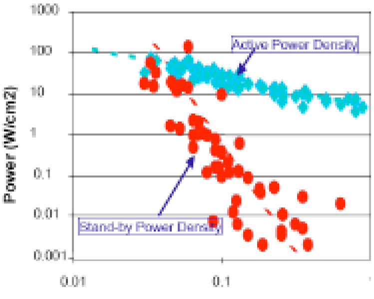


Lo et al.,IEEE EDL, May 1997

S/D Leakage



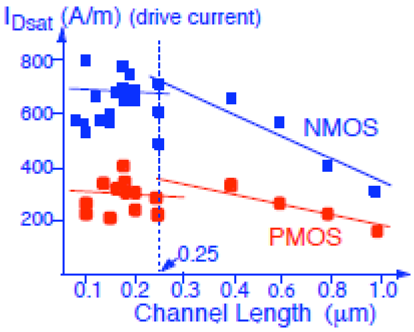
Total Power Trends



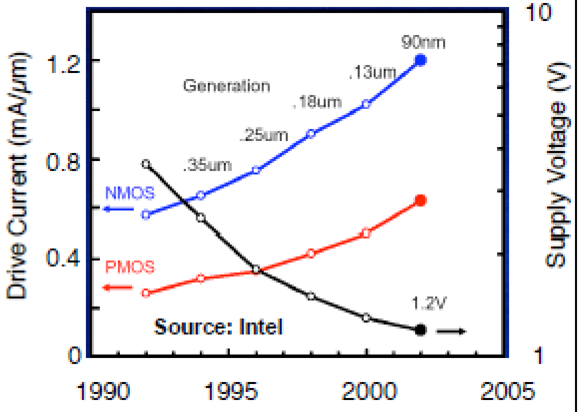
- Ability to control I_{off} will limit gate-length scaling
  - Thermionic emission over barrier
  - QM tunneling through barrier
  - Band-to-band tunneling from body to drain
- To suppress D/S leakage, need to use:
  - Higher body doping to reduce DIBL
    - ⇒ lower mobility, higher junction capacitance, increased junction leakage
  - Thinner gate dielectric to improve gate control ⇒ higher gate leakage
  - Ultra-shallow S/D junctions to reduce DIBL ⇒ higher R_{series}

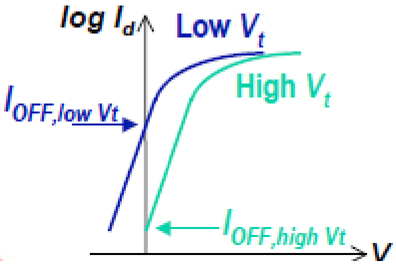
MOSFET Scaling Problem: Saturation of I_{Dsat}

Constant OFF current Limit



Relaxed OFF current Limit

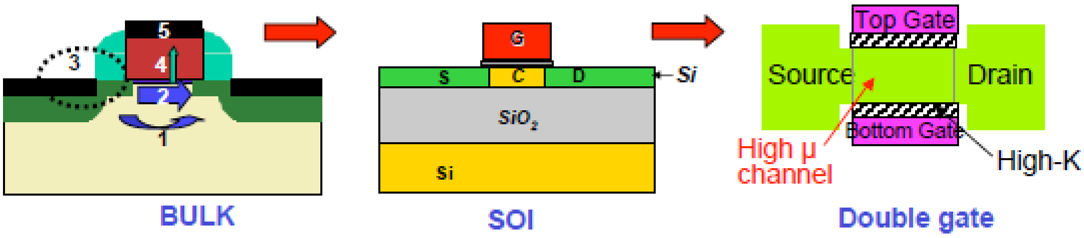




- Data from IBM, TI, Intel, AMD, Motorola and Lucent
- Low OFF current desirable

Changhoon Choi, PhD Thesis, Stanford Univ., 2002

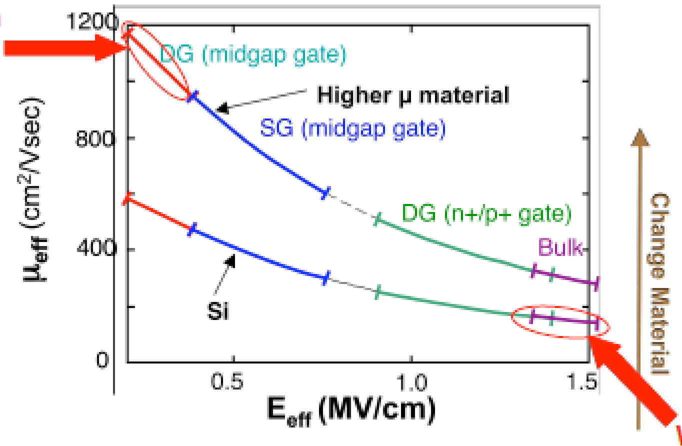
New Structures and Materials for Nanoscale MOSFETs



1. Electrostatics - Double Gate
   - Retain gate control over channel
   - Minimize OFF-state drain-source leakage
2. Transport - High Mobility Channel
   - High mobility/injection velocity
   - High drive current for low intrinsic delay
3. Parasitics - Schottky S/D
   - Reduced extrinsic resistance
4. Gate leakage - High-K dielectrics
   - Reduced power consumption
5. Gate depletion - Metal gate

Combining New Device Structures with New Materials

We will be here with these innovations

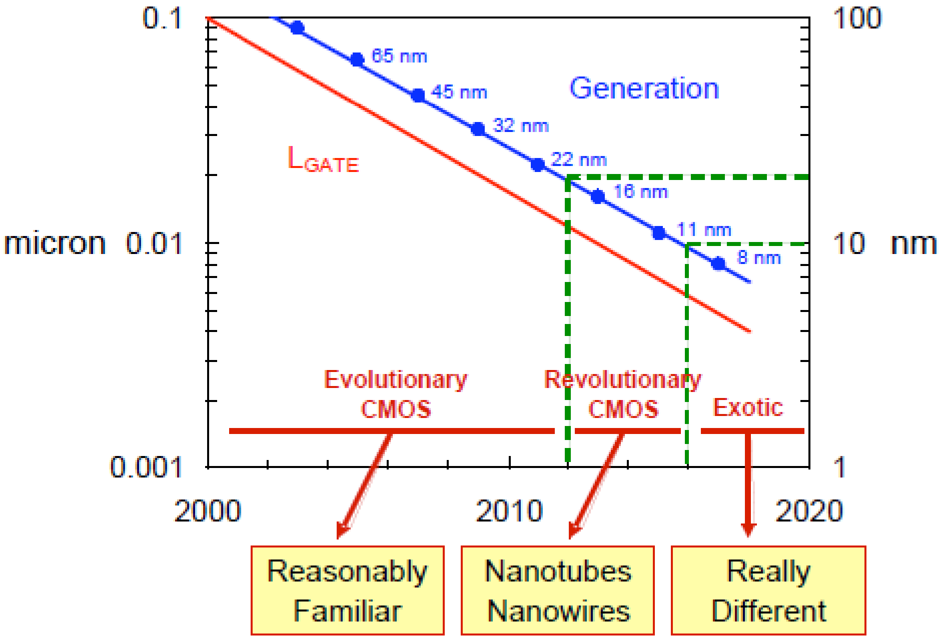


We are here today

Change Architecture allows reduced doping

- With better injection and transport we may be able to improve MOSFET I_{ON}
- With better electrostatics we may be able to minimize I_{off}

Nanotechnology Eras

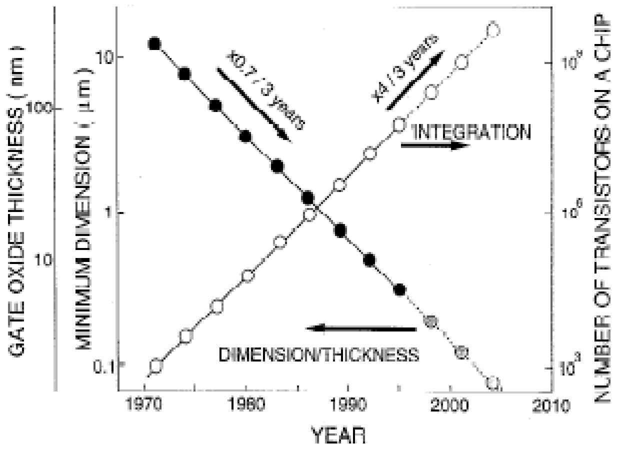


Reasonably Familiar

Nanotubes Nanowires

Really Different

Scaling of MOS Gate Dielectric



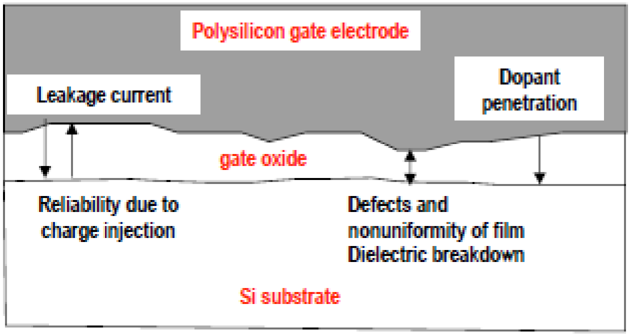
$$I_D \propto g_m \propto \frac{K}{thickness}$$

(Ref: S. Asai, Microelectronics Engg., Sept. 1996)

Gate SiO_2 thickness is approaching < 10 Å to improve device performance

- **How far can we push MOS gate dielectric thickness?**
- **How will we grow such a thin layer uniformly?**
- **How long will such a thin dielectric live under electrical stress?**
- **How can we improve the endurance of the dielectric?**

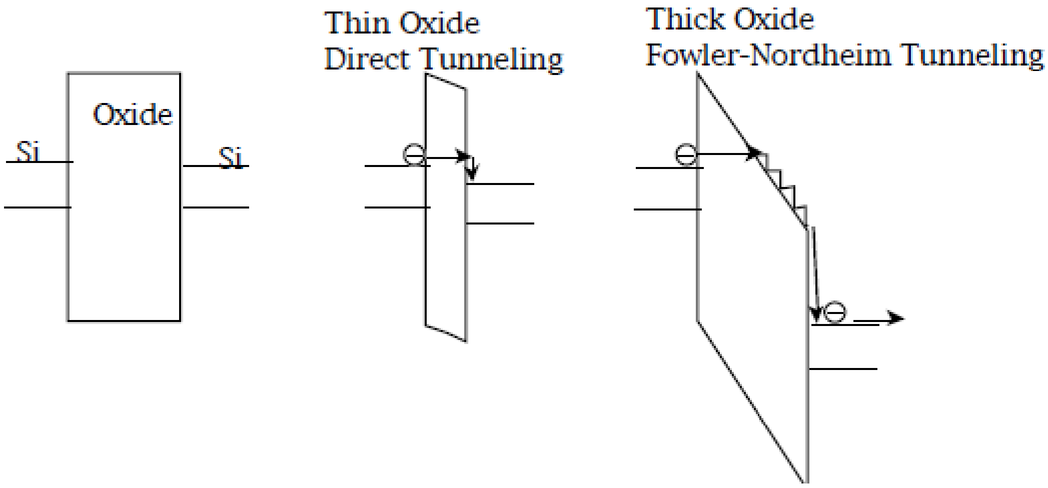
Problems in Scaling of Gate Oxide



- Below 20 Å problems with SiO_2
  - – Gate leakage => circuit instability, power dissipation
  - – Degradation and breakdown
  - – Dopant penetration through gate oxide
  - – Defects

Problems caused by conduction in ultrathin gate oxide

As we decrease the gate dielectric thickness, the conduction through the dielectric film becomes appreciable. This may increase power dissipation and cause problems for circuit stability. Increased leakage due to direct tunneling through the gate dielectric may make dynamic and static circuits unstable.

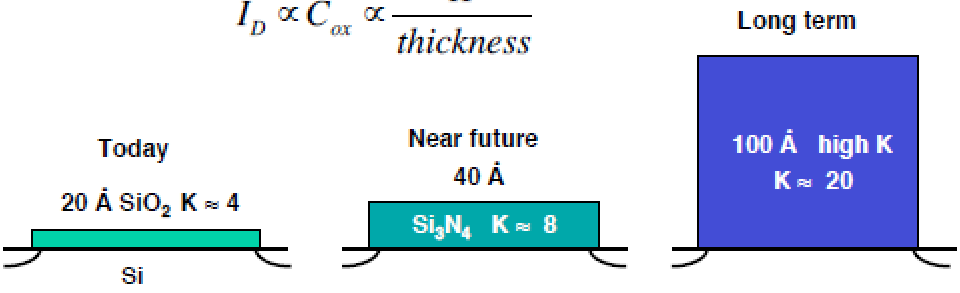


High-k MOS Gate Dielectrics

$$I_{channel} \propto \text{charge x source injection velocity}$$
$$\propto \text{(gate oxide cap x gate overdrive)}\ v_{inj}$$
$$\propto C_{ox}\,(V_{GS} - V_T)\ E_{source}\ \mu_{inj}$$

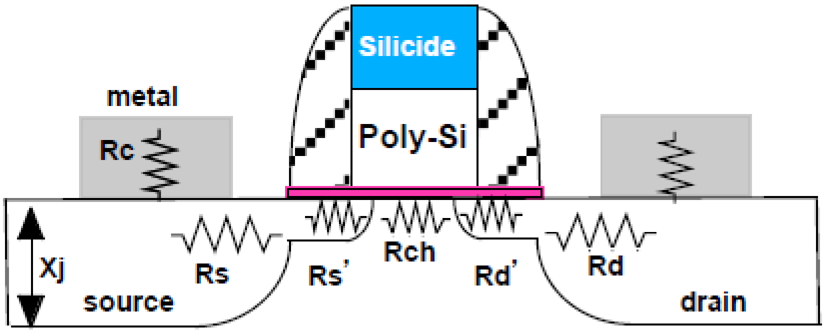
Historically C_{ox} has been increased by decreasing gate oxide thickness. It can also be increased by using a higher K dielectric

$$I_D \propto C_{ox} \propto \frac{K}{thickness}$$



Higher thickness -> reduced gate leakage

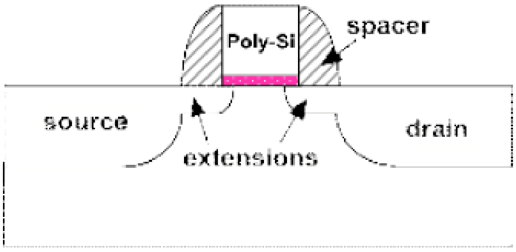
Scaling of Ohmic Contacts and Junctions



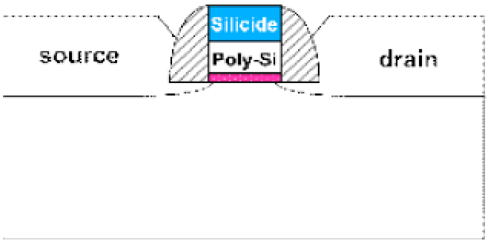
- Device scaling dictates shallow junctions.
- How will we form such shallow junctions?
- How will we make low resistance contacts to them?
- What will be the impact of the resistance of the contacts and junctions?

Solutions to Shallow Junction Problem

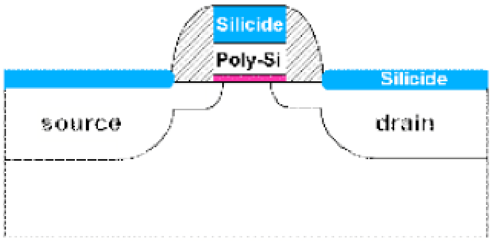
Shallow extension implants to minimize (DIBL)



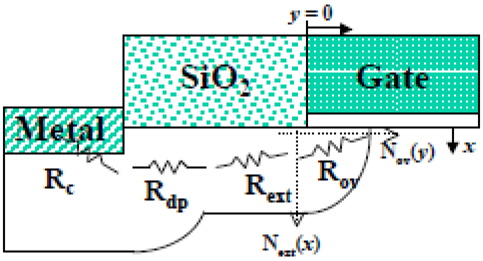
Elevated source/ drain to minimize (DIBL)

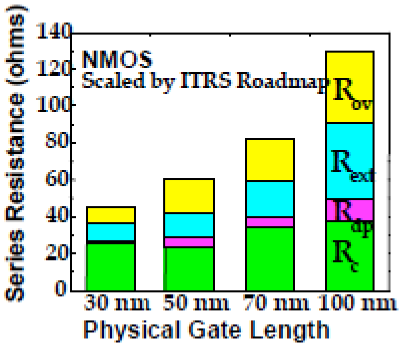


Silicidation to junction minimize resistance



Source/Drain Resistance





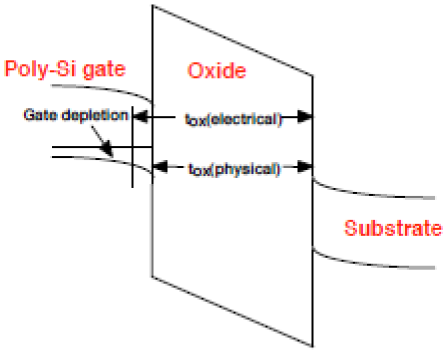
Source: Jasonn Woo, UCLA

Problem in junction scaling:

- Sheet resistance of a junction is a strong function of doping density
- Maximum doping density is limited by solid solubility and it does not scale
- Silicidation can minimize the impact of junction sheet resistance (R_s,R_d)
- Contact resistance R_c is one of the dominant components for future technology

Problems with Poly-Si Gate.

This occurs because of high E - field due to a combination of higher supply voltage and thinner gate oxide.

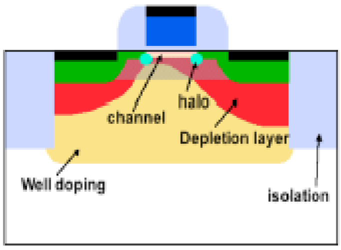


- Effect of depletion is to increase effective t_{ox} and thus reduce C_{ox}
- A reduced C_{ox} implies reduction in g_m and thus I_D(on)
- Ionized impurities in the gate electrode cause "remote charge scattering"
  - ⇒ Reduced mobility

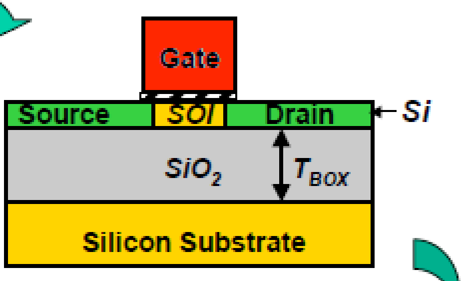
Need metal gate electrode with proper workfunction

Evolution of MOSFET Structures

BULK



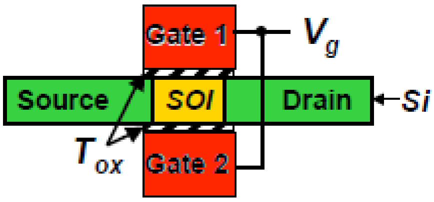
Ultra-Thin Body Single Gate SOI



Advantages of Ultra-Thin Body SOI

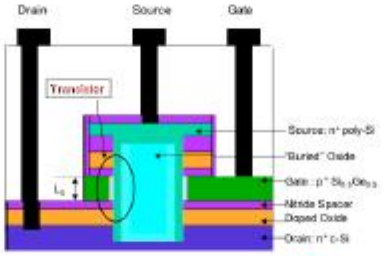
- Depleted channel ⇒ no conduction path is far from the gate
- Short channel effects controlled by geometry
- Steeper subthreshold slope
- Lower or no channel doping
- Higher mobility
- Reduced dopant fluctuation

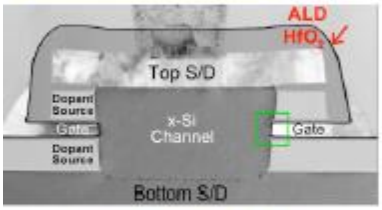
Ultra-Thin Body Double Gate SOI



Non Planar MOSFETs

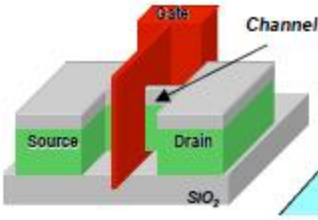
Vertical FET

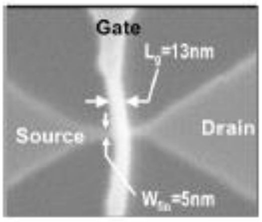




Stanford, AT&T

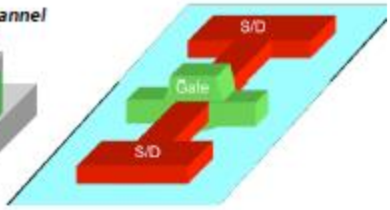
Double Gate FinFET

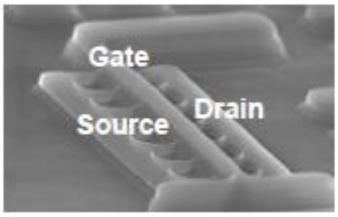




UC Berkeley

Tri Gate FET

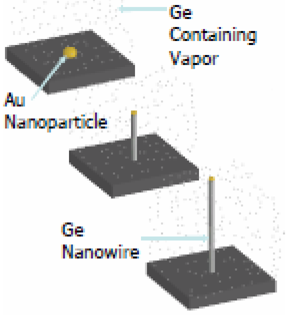




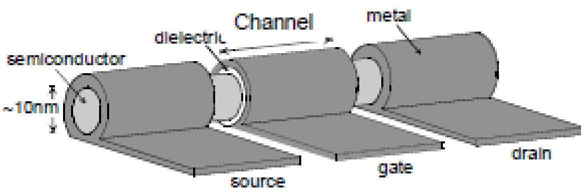
Intel

Nanowire and Nanotube FETs

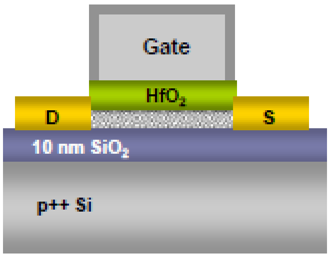
Ge NW Growth



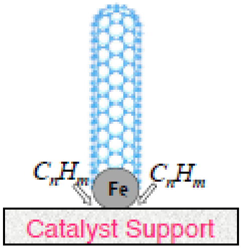
ALD HfO_2 Coated of Ge NW FET



Carbob Nanotube MOSFET

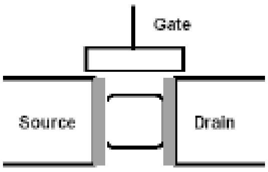


Carbob Nanotube Growth



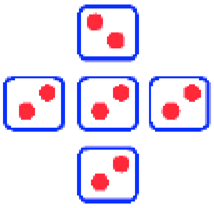
Key Challenge: Controlled growth

Seemingly Useful Devices



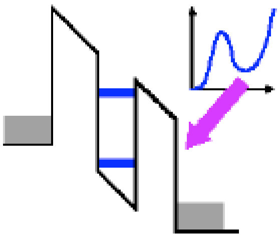
Single Electron Transistors (SET)

Limited Current Drive
Cryogenic operation



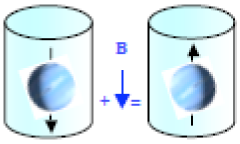
Quantum Dot

Limited Fan-Out
Critical dimension control



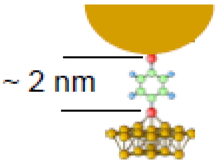
Resonant Tunneling Diode

Challenging fabrication
and process integration



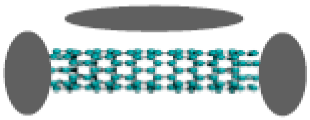
Spintronics

Need high spin injection
and long spin coherence time



Molecular Device

Limited thermal stability
New architectures needed



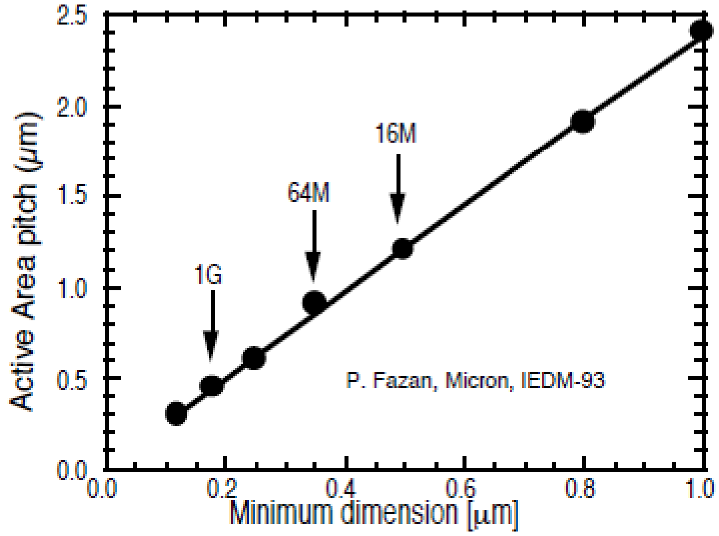
Carbon Nanotubes

Controlled growth

- In general this device scaling methodology does not take into account many other chip performance and reliability issues, e.g., interconnects, contacts, isolation, etc.

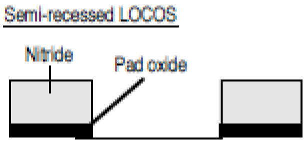
- These factors are now becoming an obstacle in the evolution of integrated circuits.

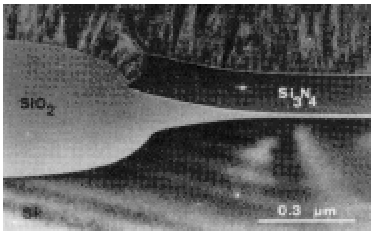
Device Isolation pitch as a function of minimum dimension



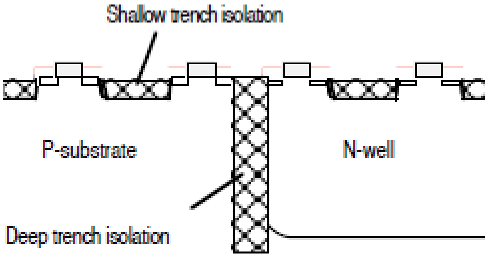
With decreasing feature size the requirement on allowed isolation area becomes stringent.

Scaling of Device Isolation



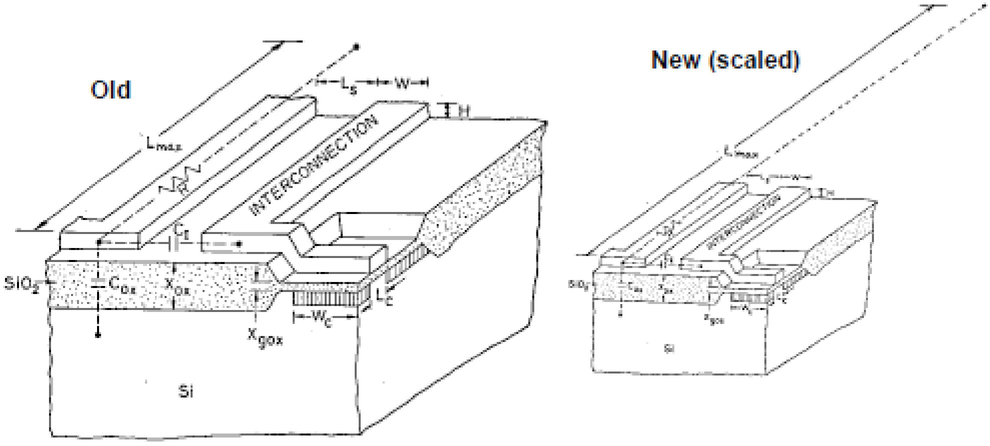


LOCOS based isolation technologies have serious problems in loss of area due to bird's beak.

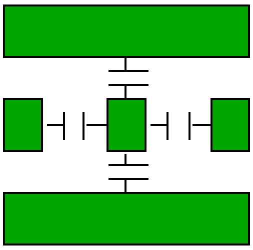


Trench isolation can minimize area loss

Scaling of interconnections

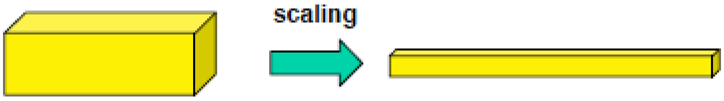


- Bigger chip => longer interconnects
- Scaling to smaller dimensions => reduced cross section
- Larger R, L and C

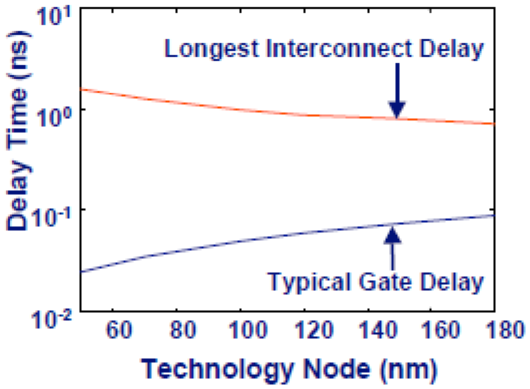


Higher Packing Density
⇓
Decreased Space Between Interconnects
⇓
Higher RC-Delay

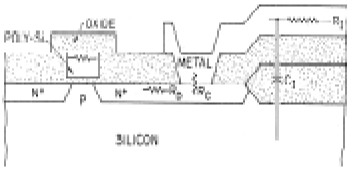
Interconnect Delay Is Increasing



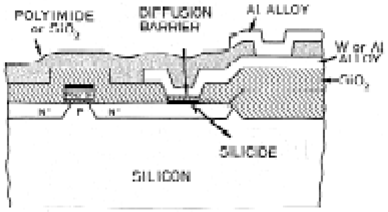
- Chip size is continually increasing due to increasing complexity
  - Increase in R, L and C
- Device performance is improving but interconnect delay is increasing
- Need better materials
  - Metal with lower resistivity
  - Dielectrics with lower K



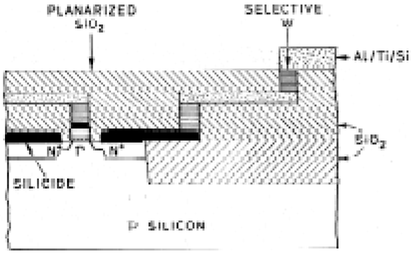
Advances in Backend Technology



1970's Poly-Si gate
Aluminum

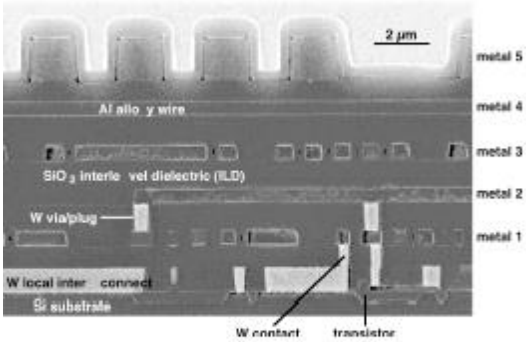


1980's Aluminum alloys
Silicide contacts
Polycide gates
Local planarization

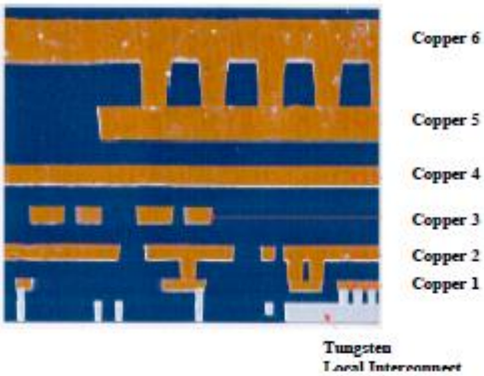


1990's Layerd aluminum/titanium
Salicides
CVD tungsten plugs
Shallow trench isolation
Global planarization

Current Interconnect Technologies

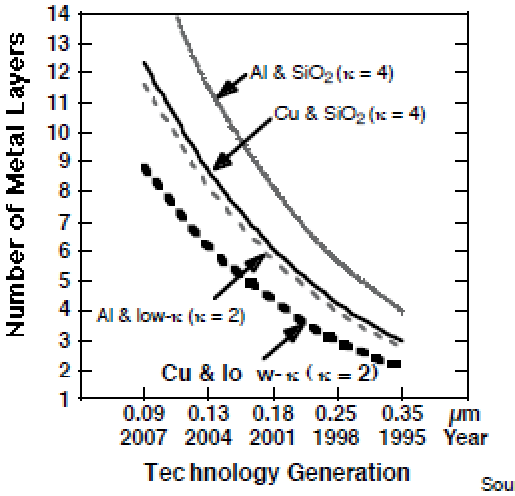


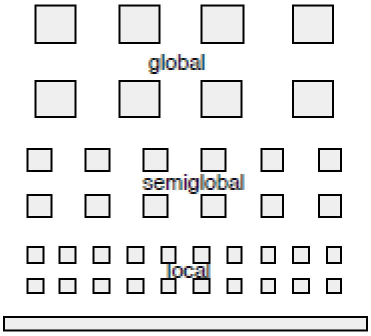
Current Al technology
(Courtesy of Motorola)



Current Cu technology
(Courtesy of IBM)

Why Cu and Low-k Dielectrics?



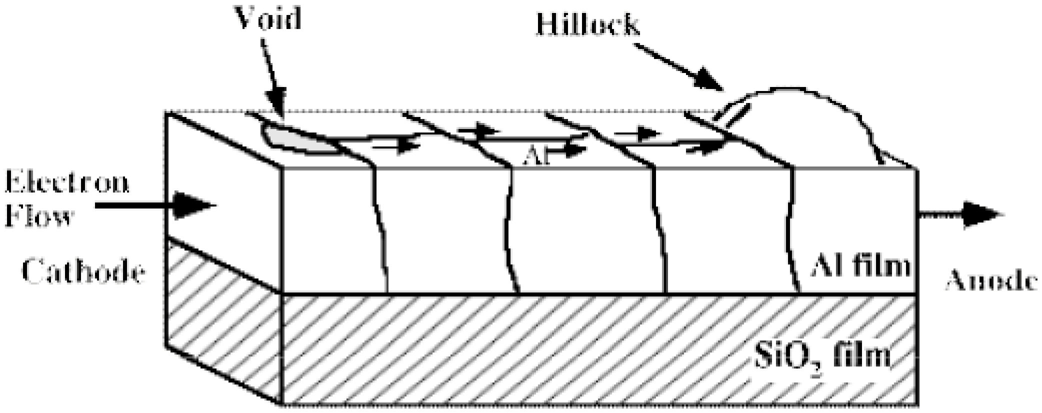


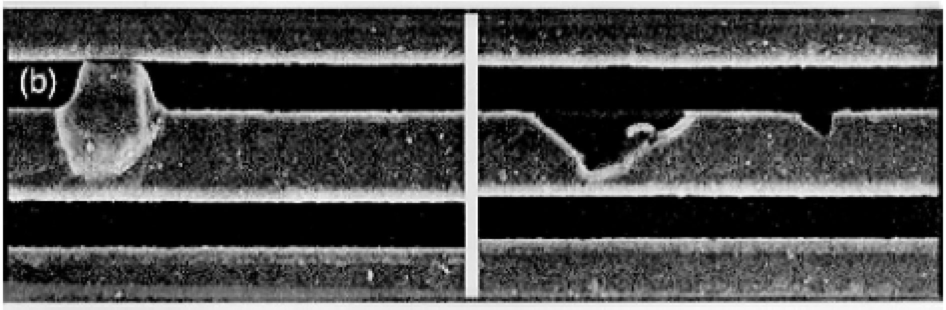
Source: Y.Nishi

Reduced resistivity and dielectric constant results in reduction in number of metal layers as more wires can by placed in lower levels of metal layers.

Electromigration

Electromigration due to electron wind induced diffusion of Al through grain boundaries





SEM of hillock and voids formation due to electromigration in an Al(Cu,Si) line

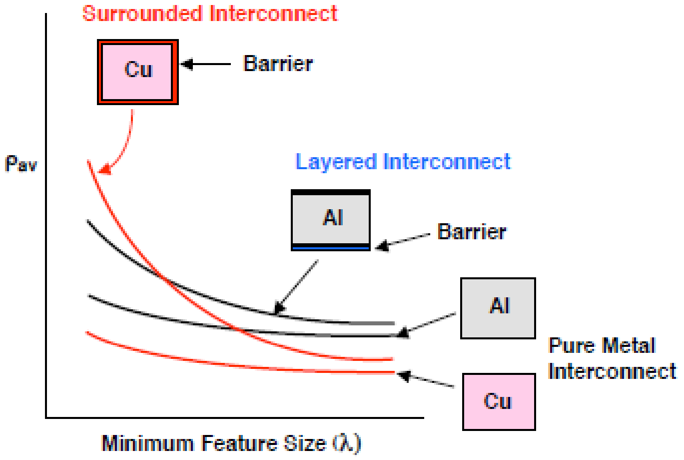
Mean time to failure due to electromigration is given by

$$MTF = \frac{A}{r^m J^n} \exp\left(\frac{E_a}{kT}\right)$$

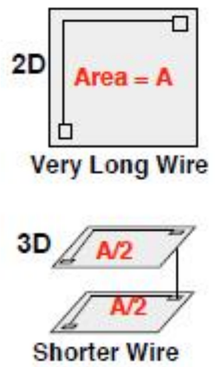
Problems in Scaling of Interconnections

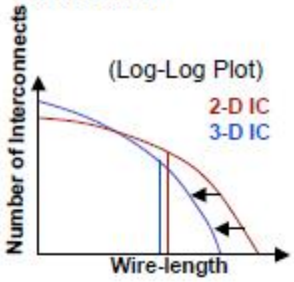
AS λ DECREASES

- Resistivity increases as grain size decreases
- Resistivity increases as main conductor size decreases but not the surroundingbarrier size



3-D Integration: Motivation

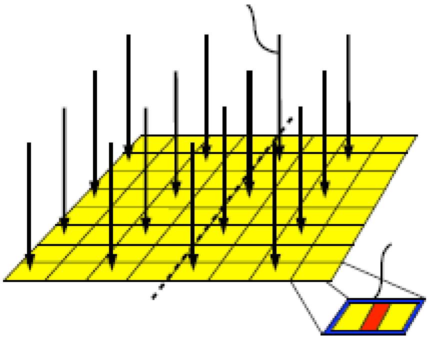




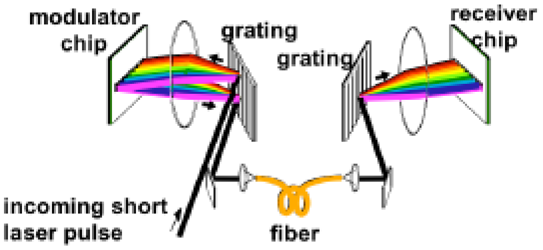
- Integration of heterogeneous technologies possible, e.g., memory & logic, optical I/O
- Reduce Chip footprint
- Replace long horizontal wires by short vertical wires
- Interconnect length ⇓ and therefore R, L, C ⇓
  - Power reduction
  - Delay reduction

Can Optical Interconnects help?

On-Chip Optical Interconnects



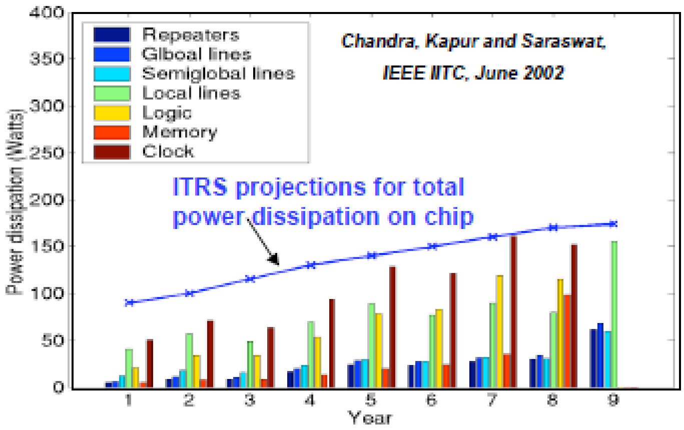
Chip-to-chip Optical Interconnects



Can potentially address many problems of Cu/low-k wires

- **On-Chip Links**
  - Reduce delay
- **Clocking and Synchronization**
  - Reduce jitter and skew
- **High Bandwidth off-chip Links**

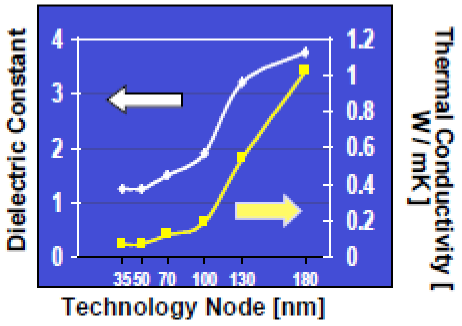
Result: scaling of power components

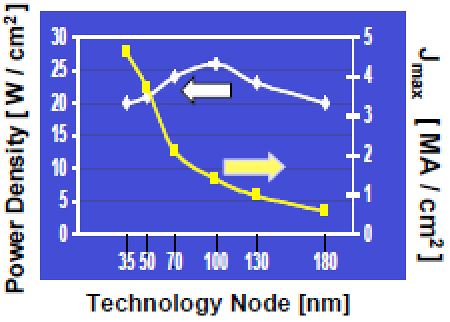


- Dynamic Power: CV^2f
- Leakage power: devices
- Short circuit power during switching
- Static power, e.g., analog components (sense amps etc.)

Power increasingly becoming the performance bottleneck for high-end microprocessors

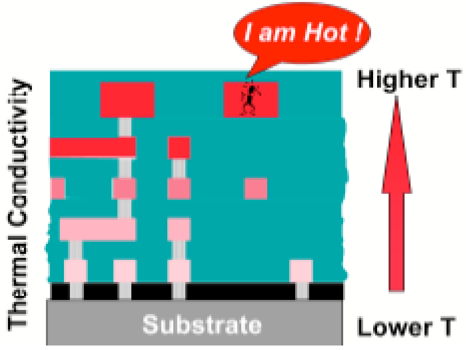
Thermal Behavior in ICs





- Thermal conductivity of low-k insulators is poor
- Thermal impedance increases
- Energy dissipated (CV^2f) is increasing as performance improves
- Average chip temperature is rising

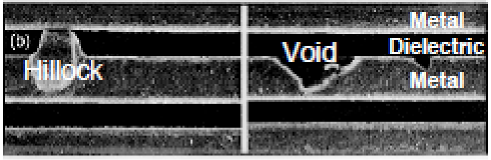
The problems Caused by Increased Power



>100A will flow on these wires

RELIABILITY

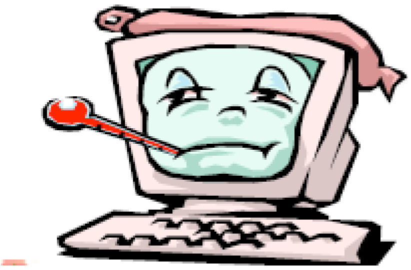
Electromigration induced hillocks and voids



Mean time to failure

$$MTF = \frac{A}{r^m J^n} \exp\left(\frac{E_a}{kT}\right)$$

10˚C ⇑ , MTF ⇓ 50%



PERFORMANCE

As T ⇑ R ⇑, RC delay ⇑
10˚C ⇑ , Speed ⇓ 5%

Milestones in Our Industry
1964 Solid Logic Tech. ………… 2014 22nm CMOS Tech.

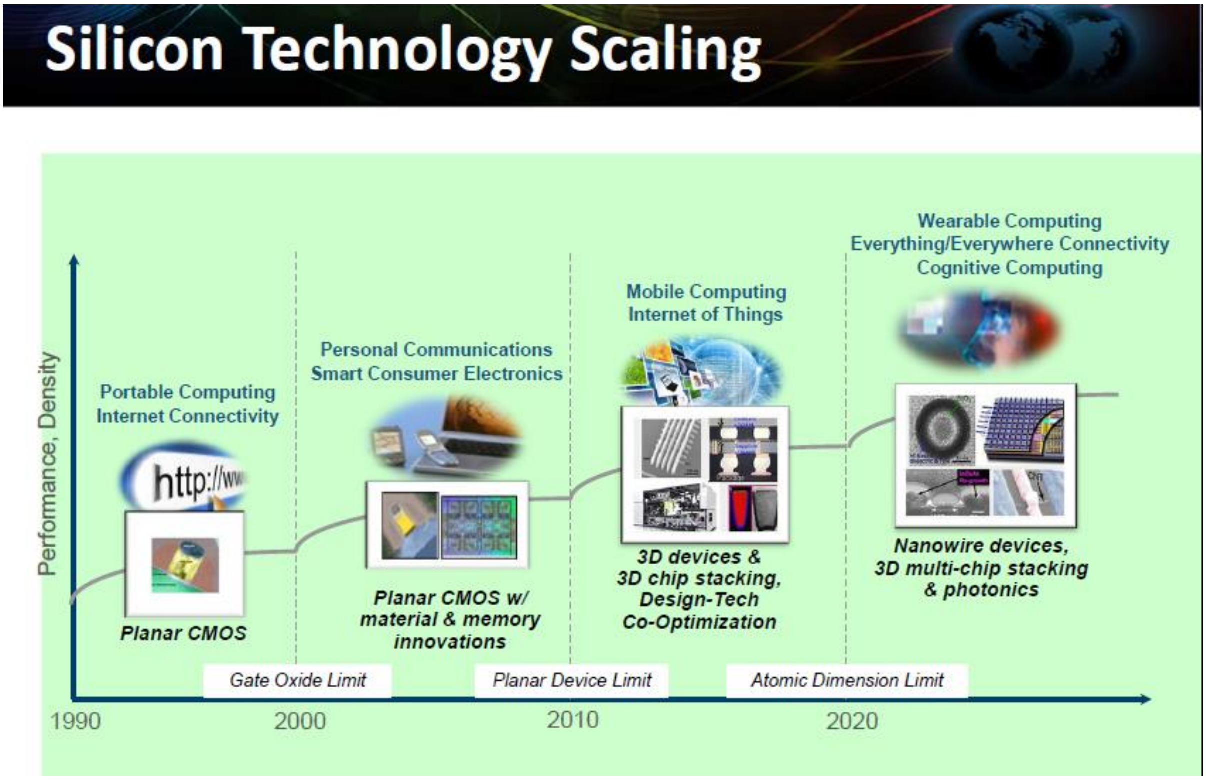
Make the Extraordinary Possible
IBM Mainframe50

IBM System 360
The machine that defined the computer industry and the modern IBM

IBM POWER8 Systems
Open Innovation for Big Data, Cloud, and Analytics

1964 - Transistor
SLT module
6 transistors, 4 resistors

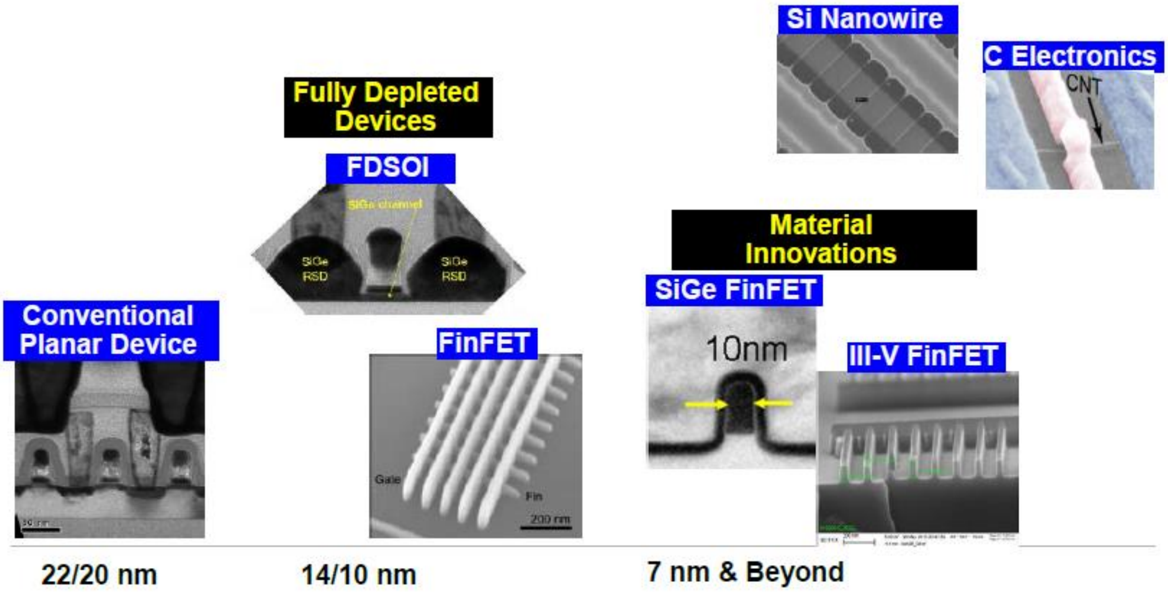
2014 – POWER8 Processor
22nm SOI eDRAM technology, 650mm^2
12 cores and 96MB of on-chip memory
4.2 billion transistors



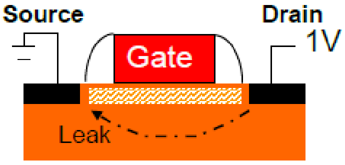
Silicon Technology Scaling

Device Research Pipeline

Scaling beyond 20nm requires alternative device structures and new material innovations.



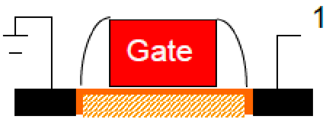
Device Innovation – Fully Depleted Devices



PDSOI

Gate controls this.

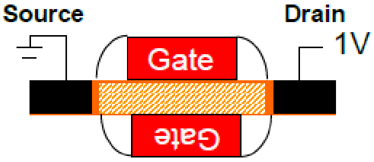
Gate can not control below that. So current can leak through there.



FDSOI

Gate controls this.

No leakage path.

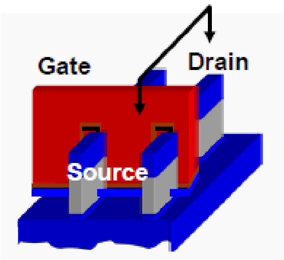


FinFET

Gates control this.

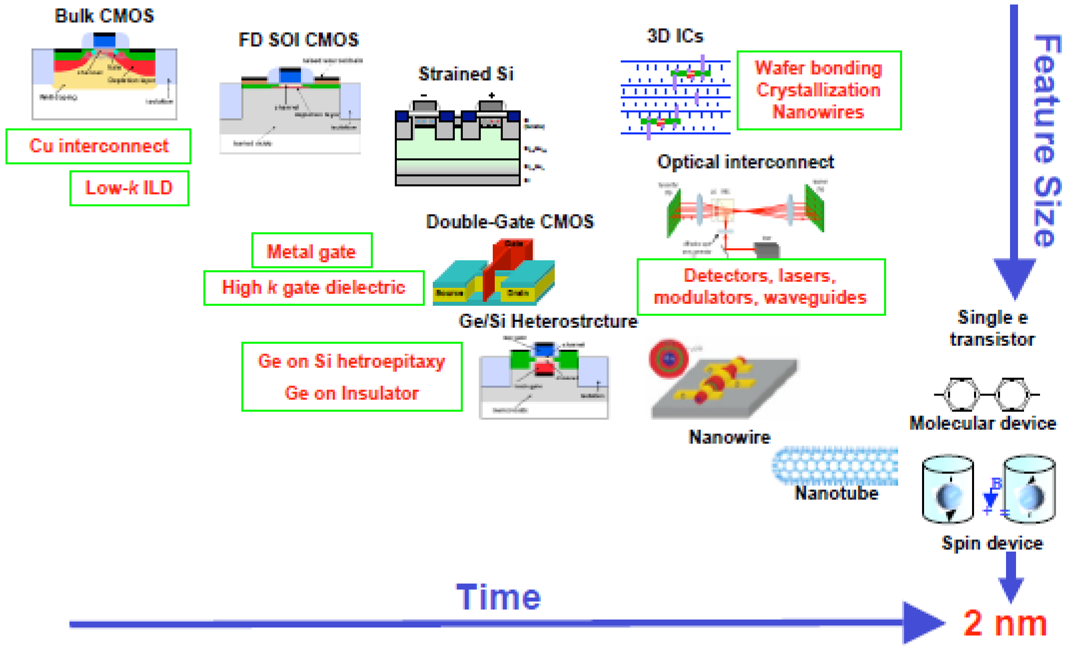
No leakage path.

Have more Si and thus can carry more current.



- **Better Electrostatics → Stronger Gate Control**
  - Lower V_t for the same leakage
  - Shorter channel for the same V_t
- **Reduced Channel Doping → Better SRAMs**
  - Less doping-driven threshold fluctuation
  - Lower supply voltage (V_{min}) – by about 150mV
  - Lower voltages means lower power – up to 40%

Conclusion: Technology Progression



Bulk CMOS

Cu interconnect

Low-*k* ILD

FD SOI CMOS

Strained Si

Metal gate

High *k* gate dielectric

Double-Gate CMOS

Ge/Si Heterostrcture

Ge on Si hetroepitaxy
Ge on Insulator

3D ICs

Wafer bonding
Crystallization
Nanowires

Optical interconnect

Detectors, lasers, modulators, waveguides

Nanowire

Nanotube

Feature Size

Single e transistor

Molecular device

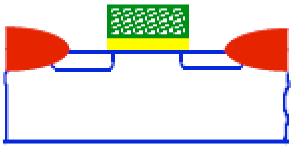
Spin device

Time

2 nm

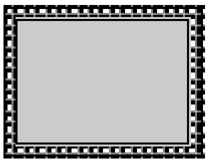
Summary

MOS Transistor in 2010



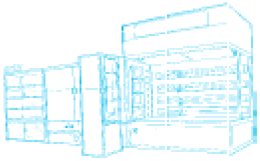
Gate oxide thickness < 1nm
Channel Length < 2nm
Junction depth < 1-2nm
Size of an atom ~ 5 Å

A Circuit in 2010



10^{10} components
Integrated digital, analog, sensors

A Factory in 2010



Approaching $10 billion

Questions we are trying to answer

- How can we continue the Moore's law
- What will be new materials, devices, circuits, sensors, equipment, simulators, etc.
- How will we design them?

Acknowledgements

1. **Dr. Gary Patton , Vice President, IBM Semiconductor**
2. **Prof. Krishna Saraswat , Stanford University**