Advanced VLSI Design (ECE 695KR)

Kaushik Roy
Professor of ECE
Purdue University

Acknowledgement

- Professor Supriyo Datta
- Swaroop Ghosh
- Professor Chris Kim
- Professor Swarup Bhunia
- Professor Saibal Mukhopadhyay
- SRC, Intel, AMD,..

Course Overview

- **Targeted for graduate students who have already taken basic VLSI design classes**
- **Real world challenges and solutions in designing high-performance and low-power circuits**
- **Relations to VLSI Design**
  - » **Recent developments in digital IC design**
  - » **Project oriented**
  - » **Student participation: class presentation**

Prerequisite

- **MOS VLSI Design or equivalent**
  - » **MOS transistor**
  - » **Static, dynamic logic**
  - » **Adders, Multipliers,..**
  - » **….**
- **Familiarity with VLSI CAD tools**
  - » **Magic or Cadence: LVS, DRC**
  - » **HSPICE**
- **Basic knowledge on solid-state physics**

Class Materials

- **Lecture notes: primary reference**
- **K. Roy, S. Prasad, *Low Power CMOS VLSI Circuit Design*, John Wiley**
- **A. Chandrakasan, W. Bowhill, F. Fox, *Design of High-Performance Microprocessor Circuits*, IEEE Press, 2001.**
- **Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices*, Cambridge University Press, 2002.**
- **J. Rabaey, A. Chandrakasan, B. Nikolic, *Digital Integrated Circuits: A Design Perspective*, Prentice Hall, 2nd edition, 2003. (prerequisite)**

Approximate Course Outline

- Introduction -- Importance of Power, Speed, & Reliability in scaled technologies and technology requirements
  - » Digital computing and transistor as a switch. Basic switch characteristics required for efficient and low-energy computing
  - » Ion/Ioff ratio and CV/I as a measure of device "good-ness"
  - » Power dissipation and the minimum energy required for computing – comparison with today's power dissipation and computational needs. What can possibly be done to improve power while being speed-efficient.
- Transistors and scaling of technology
  - » Bottom-up approach starting from molecules to bulk
  - » Bulk-Si, Double-gate, Tri-gate devices and possible circuit implications
  - » Emerging technologies and possible circuit implications
- Device/Circuit co-design for speed/power/reliability
  - » Inverters, Logic
  - » Memory bit-cell

Course Outline (cont'd)

- Leakage components, estimation and leakage tolerant design techniques
  - » Subthreshold, Gate, Junction tunneling, GIDL, Diode leakage etc.
  - » Memory and logic design for leakage tolerance
  - » Scalability of leakage tolerant design techniques
- Parameter variations – spatial and temporal – due to process, NBTI, Hot-electron, TDDB
  - » OPC
  - » Sensors to detect process corners/reliability degradation
  - » Failure analysis for memories and logic
- Design with unreliable components
  - » Logic and memories
- Low-voltage design including digital subthreshold operations
- Interconnects and its impact on scaled technologies. Possible solutions at the technology, circuit, architecture level
- Low-power and process-tolerant design solutions at the algorithm and architecture level
  - » Digital signal processing applications
  - » General purpose computing

Grading etc.

- Semester long project (report due on the last week of class) – 60% of the grade
  - » Project of your choice but discuss with me to get approval
  - » If you have difficulty finding a topic, I can help
  - » Presentation of your work – mid-semester and end-of the semester
  - » Project should be of publishable quality and the report should be of normal conference paper format (6 pages, double column etc.)
  - » Best and second best project (determined by a panel of judges) will be rewarded with a cache prize (~$350-500/per student) and a plaque. Funded by AMD. So do your best.
- Homeworks and presentations on topics of current interest – 40% of grade. Problems will be given throughout the semester. Each student will take turns to post the homework solutions.

CAD Tools

- **Cadence**
  - » **Schematic editor, layout editor, DRC, LVS**
- **HSPICE, awaves**
- **Technology files**
  - » **TSMC 0.18µm, BPTM 70nm, …**
- **Synopsys design compiler, library compiler**
- **Taurus-device, Taurus-Medici**

- **Everyone should have some experience with these tools**

Project Topic

- **Students pick the research topic they want to work on**
- **After the literature survey, choose a paper that you would like to evaluate yourself**
- **Has to be on digital VLSI circuit/system**
  - » **Op-amp design alone is not acceptable**
  - » **Op-amp design for digital applications is acceptable**
- **Show your work's claim using your own simulations**
- **Your contribution must be clearly shown at the end**
  - » **Improve previous design**
  - » **New circuit/system, modeling technique**
  - » **Show limitation of previous techniques**
- **Talk to the instructor in case you need help**

How to Find a Project Topic?

- **Conferences**
  - » **International Solid-State Circuits Conference (ISSCC, top conference!): slides posted on IEEExplore**
  - » **Symposium on VLSI Circuits (VLSIC), DAC, ICCAD**
  - » **Custom Integrated Circuits Conference (CICC)**
  - » **……**
- **Journal**
  - » **IEEE TVLSI, IEEE TCAD, IEEE TED**
  - » **IEEE Journal of Solid-State Circuits (JSSC)**
  - » **Intel Technology Journal**
  - » **IBM Journal on R & D**
  - » **…..**

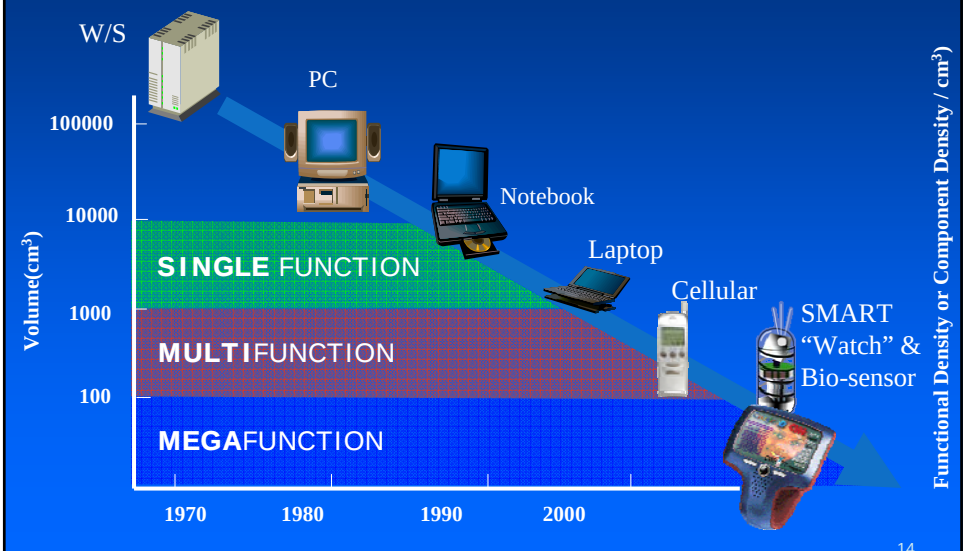
How to Find a Project Topic?

- **Funding agencies**
  - » **Research needs document (www.src.org)**
- **Presentation**
  - » **University of Michigan VLSI seminar series (www.eecs.umich.edu/vlsi_seminar/)**
  - » **Design automation conference (www.dac.com)**
- **Pick a recent issue in VLSI design (< 5 years)**
- **I suggest you start doing the literature survey ASAP**

Academic Misconduct

- **Students caught engaging in an academically dishonest practice will receive a failing grade for the course.**
- **University policy on academic dishonesty will be followed strictly.**

Trend to Miniaturization of Convergent Systems



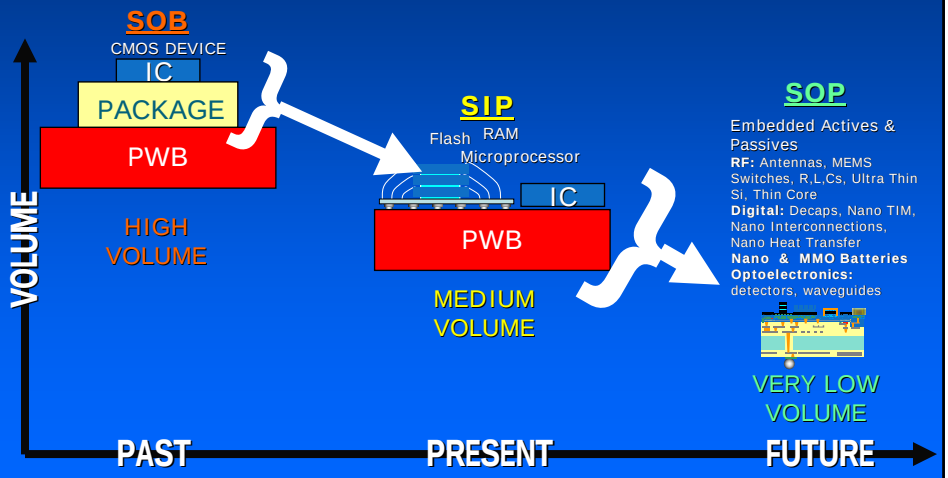
Vision of Digital Convergence by 1000X (Source: GeorgiaTech PRC)

Digital + Analog + RF + Optical+Sensors

- Computing/Internet
- Digital Audio
- Digital Imaging/Video
- Cellular/Wireless
- GPS/Satellite
- Sensors
- And, of course, timekeeping!

Package & IC Integration to 3D SOP

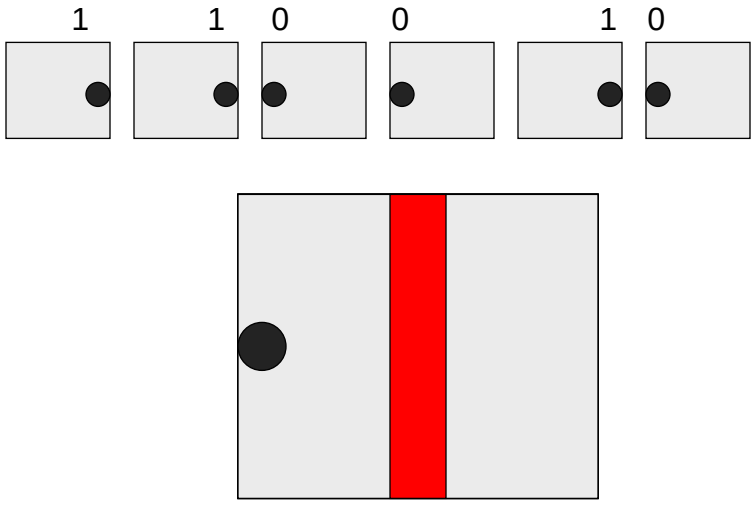
3D Metric: Transistors / cm³ or Components/cm³



A physical system as a computing medium

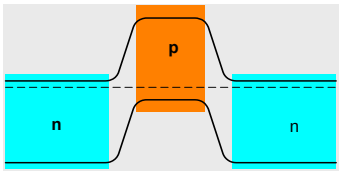
- We need to create a bit first. Information processing always requires physical carrier, which are material particles.
- *First* requirement to physical realization of a bit implies creating *distinguishable* states within a system of such material particles.
- *The second* requirement is *conditional* change of state.
- The properties of *distinguishability* and *conditional change of state* are two fundamental properties of a material subsystem to represent information. These properties can be obtained by creating *energy barriers* in a material system.

Particle Location is an Indicator of State

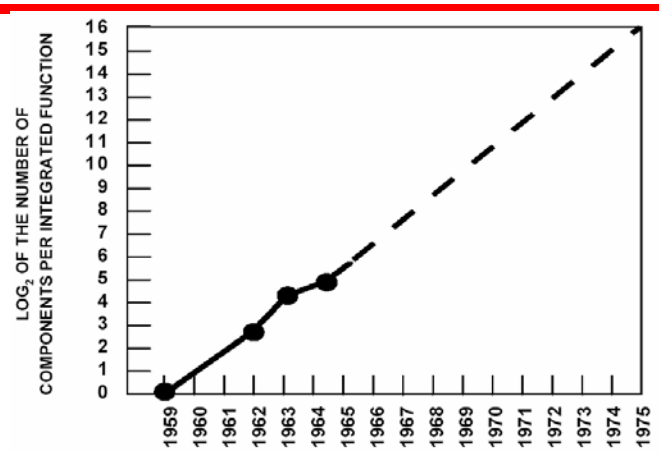


Barrier engineering in semiconductors

By doping, it is possible to create a built-in field and energy barriers of controllable height and length within semiconductor. It allows one to achieve conditional complex electron transport between different energy states inside semiconductors that is needed in the physical realization of devices for information processing.

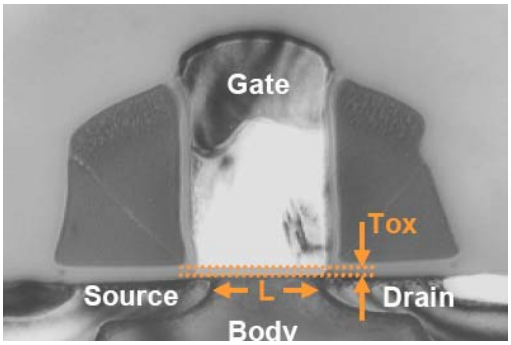


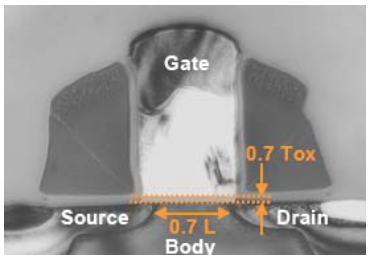
Moore's Law



- **Intel founder and chairman Gordon Moore predicted in 1965 that the number of transistors on a chip will double every 18-24 months**

Transistor Scaling





- **Constant E-field scaling: voltage and dimensions (both horizontal and vertical) are scaled by the same factor k, (~1.4), such that the electrical field remains unchanged.**

Technology Scaling

$$Dimensions \xrightarrow{scale} 0.7,\ V_{dd} \xrightarrow{scales} \beta,\ V_t \xrightarrow{scales} \beta$$

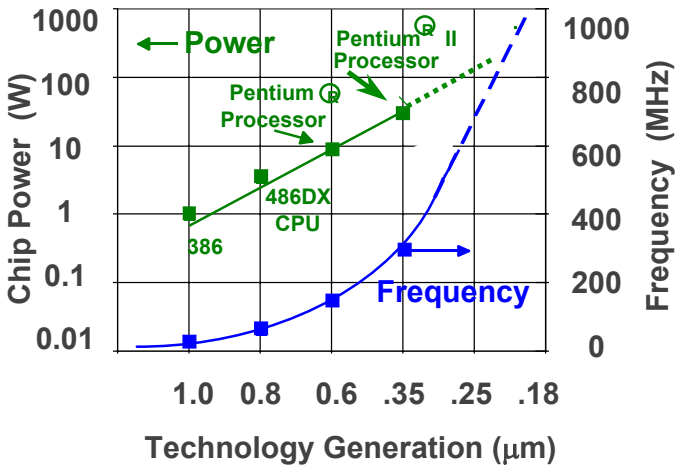
$$I = \frac{kW}{T_{ox}}(V_{dd} - V_t) \xrightarrow{scales} \frac{0.7}{0.7} \times \beta = \beta$$

$$D = \frac{CV_{dd}}{I} \xrightarrow{scales} \frac{0.7 \times \beta}{\beta} = 0.7 \quad \text{(30\% delay reduction)}$$

$$E = CV_{dd}^{\ 2} \xrightarrow{scales} 0.7\beta^{\ 2}$$

IC Frequency & Power Trends

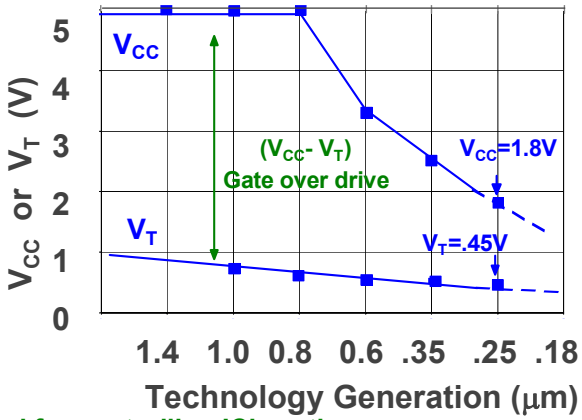
- Clock frequency improves 50%
- Gate delay improves ~30%
- Power increases 50%
- Power = $C_L V^2 f$



↑ Active switched capacitance "C_L" is increasing.

Vdd vs. Vt scaling

- Recently: constant e-field scaling, aka voltage scaling
- V_{CC} ⌫ 1V
- V_{CC} & modest V_T scaling
- Loss in gate overdrive (V_{CC}-V_T)



↑ **Voltage scaling is good for controlling IC's active power, but it requires aggressive V_T scaling for high performance**

Delay

$$\tau_d = \frac{C_L V_{DD}}{I_D}$$

$$\tau_d = \frac{C_L}{(\frac{W}{2L})\mu C_{ox} V_{DD}(1-\frac{V_T}{V_{DD}})^2}$$ Long Channel MOSFET

$$\tau_d = \frac{C_L}{WC_{ox}\upsilon_{SAT}(1-\frac{V_T}{V_{DD}})}$$ Short Channel MOSFET

$$\tau = \frac{CL^{0.5}T_{ox}^{0.5}}{V_{DD}^{0.3}(0.9-V_T/V_{DD})^{1.3}}(\frac{1}{W_n}+\frac{2.2}{W_p})$$ [1]

[1] C. Hu, "Low Power Design Methodologies," Kluwer Academic Publishers, p. 25.

Performance significantly degrades when V_{DD} approaches $3V_T$.

V_T Scaling: V_T and I_{OFF} Trade-off

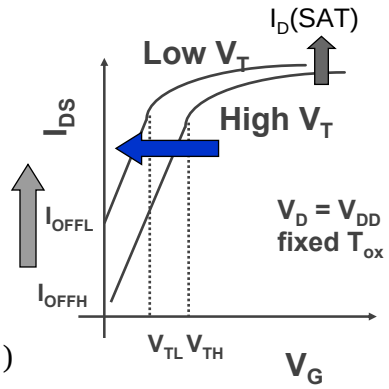
Performance vs Leakage:

$V_T \downarrow I_{OFF} \uparrow I_D(SAT) \uparrow$

$$I_{OFF} \propto I_{subth} \propto \frac{W_{eff}}{L_{eff}} K_1 e^{(V_{GS}-V_T)}$$

$$I_D(SAT) \propto \frac{W_{eff}}{L_{eff}} K_2 (V_{GS}-V_T)^2$$

$$I_D(SAT) \propto K_3 W_{eff} C_{ox} \upsilon_{SAT} (V_{GS}-V_T)$$

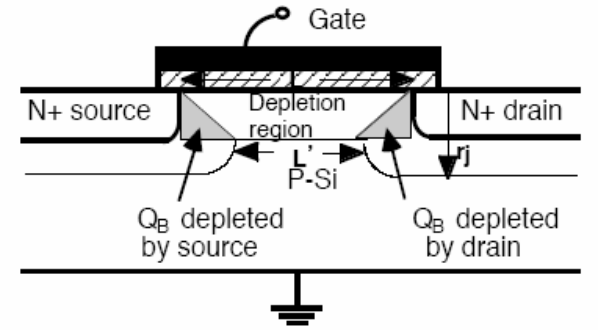


- As V_T decreases, sub-threshold leakage increases
- Leakage is a barrier to voltage scaling

Constant Field Scaling

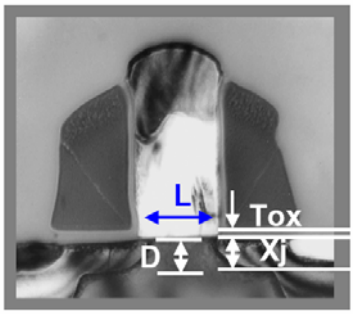
	Device and circuit parameters	Factor
Scaling assumptions	Device dimensions (t_{ox}, L, W, X_j)	1/k
	Doping concentration (N_a, N_d)	k
	Voltage (V)	1/k
Device parameters	Electric field (E)	1
	Capacitance ($C=\varepsilon A/t$)	1/k
	Current (I)	1/k
	Channel resistance (R_{ch})	1
Circuit parameters	Delay (CV/I)	1/k
	Power (VI)	$1/k^2$
	Switching energy (CV^2)	$1/k^3$
	Circuit density (1/A)	k^2
	Power density (P/A)	1

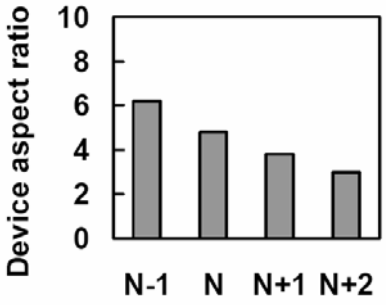
Scaling in the Vertical Dimension



- **Transistor V_t rolls off as the channel length is reduced**
- **Shallow junction depth reduces V_t roll-off**
- **However, sheet resistance increases**

Scaling in the Vertical Dimension





$$\text{Device aspect ratio} \approx \frac{L}{\sqrt[3]{T_{ox}\,\frac{\varepsilon_{si}}{\varepsilon_{ox}}\,X_j\,D}}$$

- **Vertical dimension scales less than horizontal**
- **Aggravates short channel effect (V_t roll-off)**

Constant Voltage Scaling

	Device and circuit parameters	Factor
Scaling assumptions	Device dimensions (t_{ox}, L, W, X_j)	1/k
	Doping concentration (N_a, N_d)	k
	Voltage (V)	1
Device parameters	Electric field (E)	k
	Capacitance ($C=\varepsilon A/t$)	1/k
	Current (I)	k
	Channel resistance (R_{ch})	1/k
Circuit parameters	Delay (CV/I)	$1/k^2$
	Power (VI)	k
	Switching energy (CV^2)	1/k
	Circuit density (1/A)	k^2
	Power density (P/A)	k^3

Constant Voltage Scaling

- **More aggressive scaling than constant field**
- **Limitations**
  - » **Reliability problems due to high field**
  - » **Power density increases too fast**

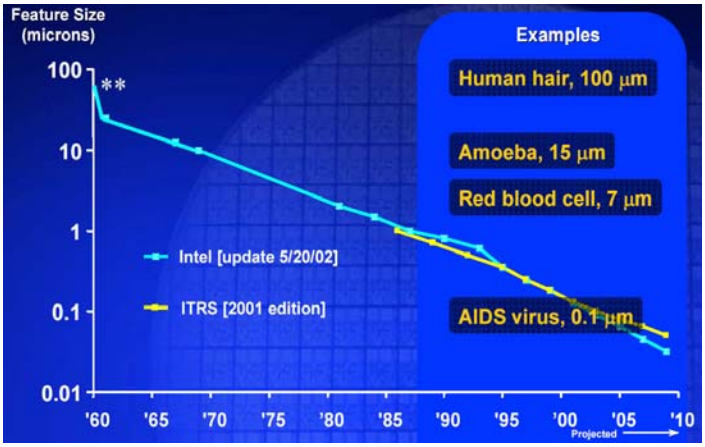
- **Both constant field and constant voltage scaling have been followed in practice**
- **Field and power density has gone up as a byproduct of high performance, but till now designers are able to handle the problems**

ITRS Roadmap

Year	2001	2003	2005	2007	2010	2013	2016
DRAM ½ pitch [nm]	130	100	80	65	45	32	22
MPU transistors/chip	97M	153M	243M	386M	773M	1.55G	3.09G
Wiring levels	8	8	10	10	10	11	11
High-perf. phys. gate [nm]	65	45	32	25	18	13	9
High-perf. VDD [V]	1.2	1.0	0.9	0.7	0.6	0.5	0.4
Local clock [GHz]	1.7	3.1	5.2	6.7	11.5	19.3	28.8
High-perf. power [W]	130	150	170	190	218	251	288

- **International Technology Roadmap for Semiconductors 2002 projection (http://public.itrs.net/)**

Transistor Scaling

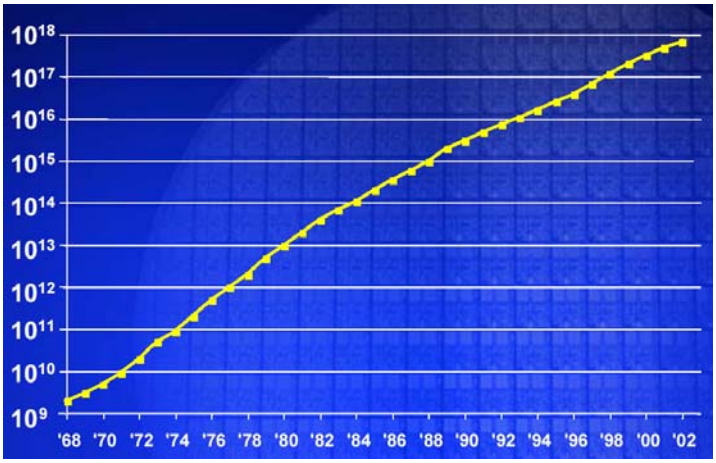


- **65nm is in production, 45-30nm in research phase**
- **New technology generation introduced every 2-3 years**

Cost per Transistor

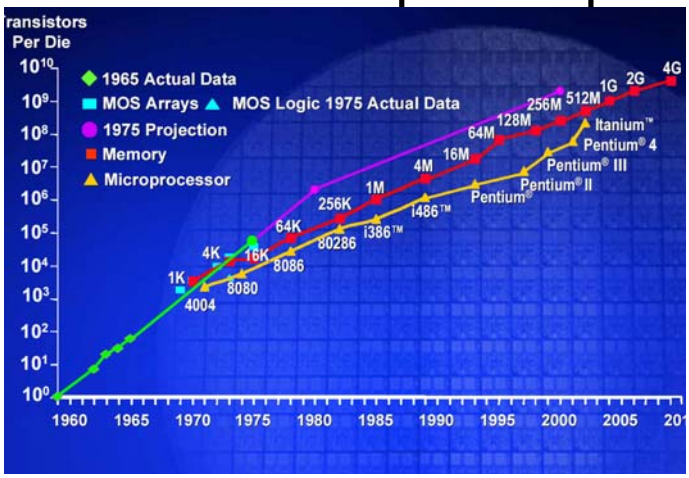
- **You can buy 10M transistors for a buck**
- **They even throw in the interconnect and package for free**

Transistors Shipped Per Year



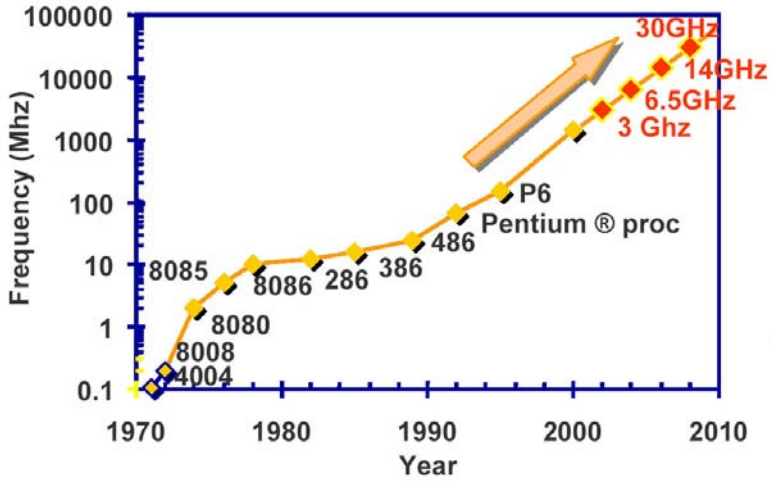
- **Today, there are about 100 transistors for every ant**
  - Gordon Moore, ISSCC '04

Transistors per Chip



- **1.7B transistors in Montecito (next generation Itanium)**
- **Most of the devices used for on-die cache memory**

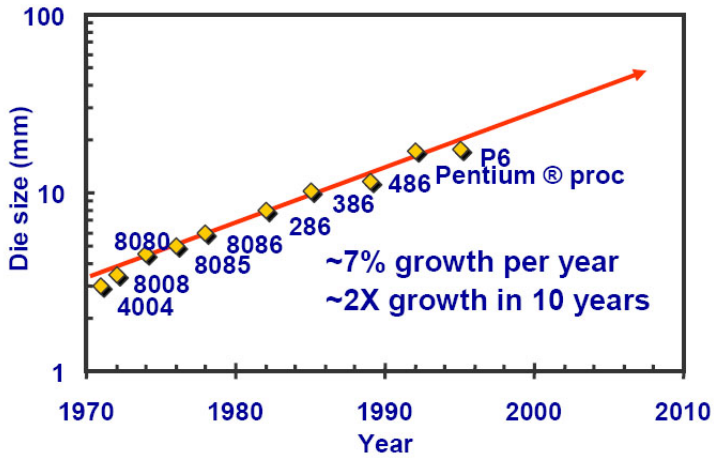
Chip Frequency



S. Borkar

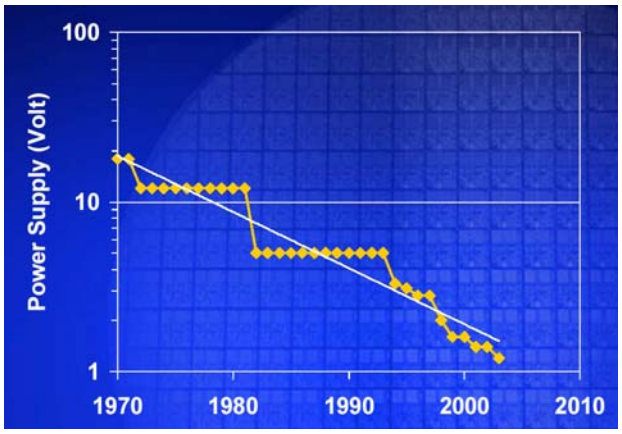
- **30% higher frequency every new generation**

Die Size



- **~15% larger die every new generation**
- **This means more than 2X increase in transistors per chip**

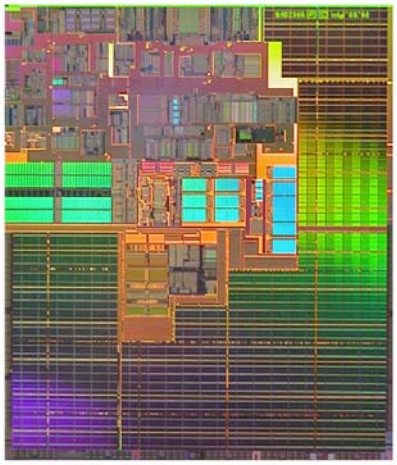
Supply Voltage Scaling



- **Supply voltage is reduced for active power control**

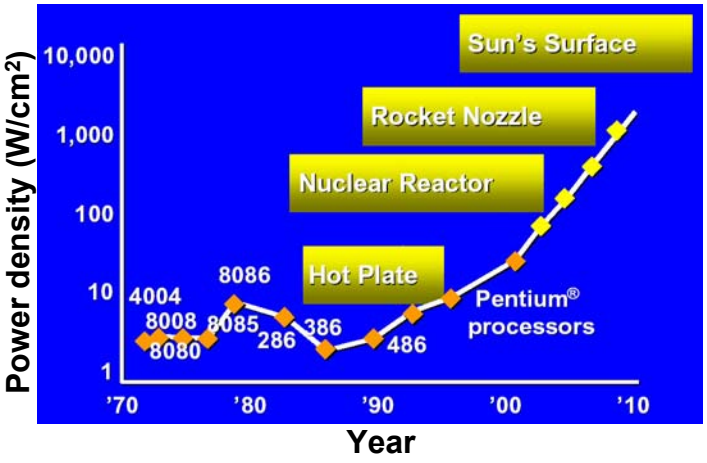
$$P_{active} \propto C V_{dd}^2 f$$

4 Decades of Transistor Scaling:
Itanium 2 Processor



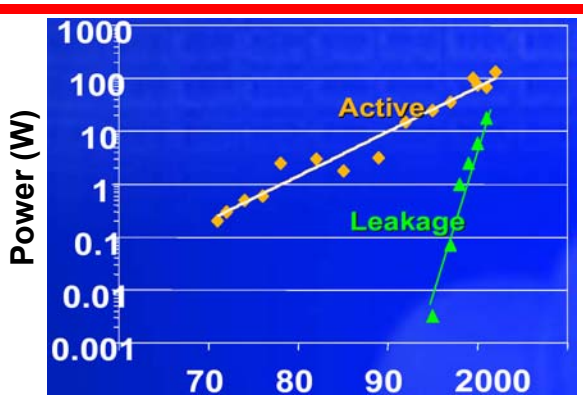
- 130nm process
- 410M transistors
- 374mm² die size
- 6MB on-die L3 cache
- 1.5GHz at 1.3V
- 6.4GB/s 400MT/s 4-way bus interface
- System compatible with existing Itanium 2 platforms
- Extensive RAS, DFT and DFM features

Power Density



- **High-end microprocessors: Packaging, cooling**
- **Mobile/handheld applications: Short battery life**

Active and Leakage Power

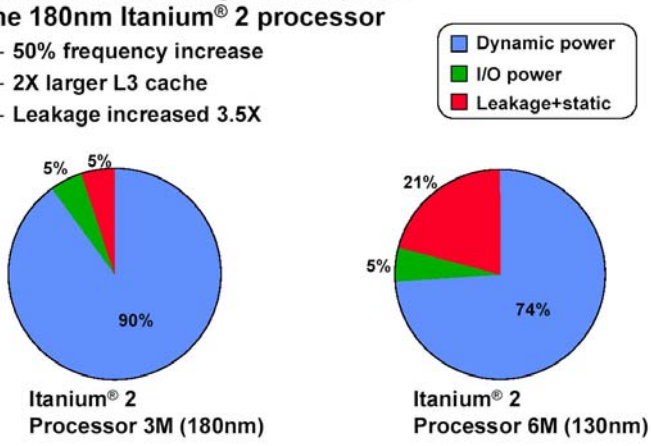


$$delay \propto \frac{C_L}{1 - \frac{V_t}{V_{dd}}} \qquad I_{leak} \propto \exp\left(\frac{-V_t}{mkT/q}\right)$$

- **Transistors are becoming dimmers**

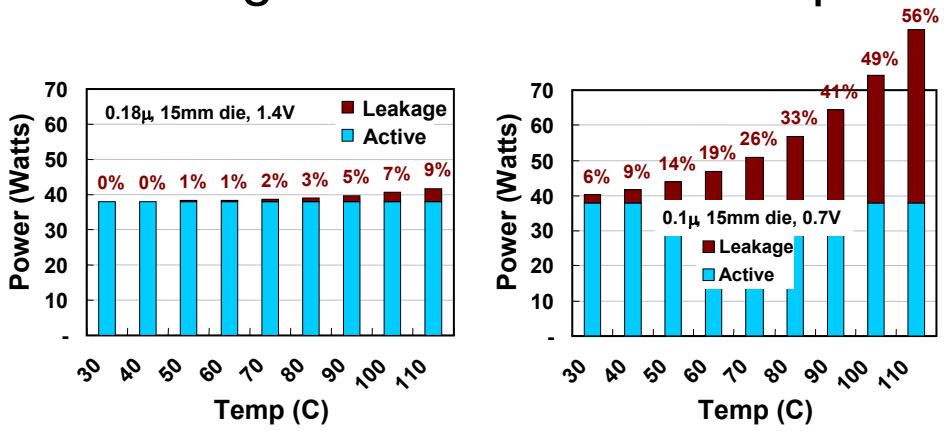
Leakage Power Crawling Up in Itanium 2

- Same thermal design envelope as the 180nm Itanium® 2 processor
  - 50% frequency increase
  - 2X larger L3 cache
  - Leakage increased 3.5X



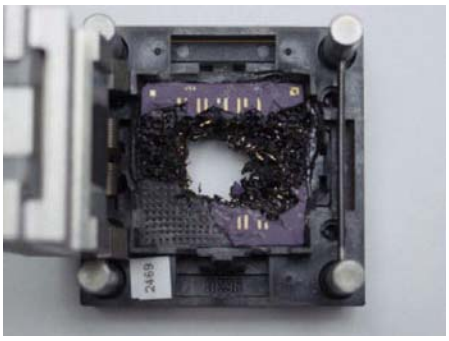
- **Transistor leakage is perhaps the biggest problem**

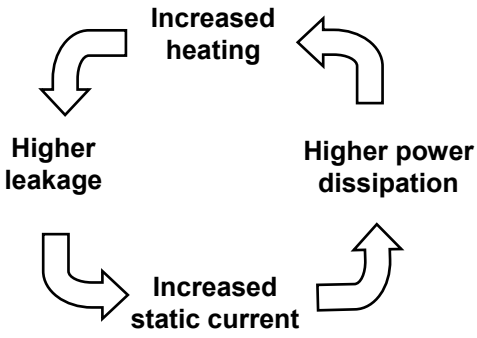
Leakage Power versus Temp.



- **Leakage power is problematic in active mode for high performance microprocessors**

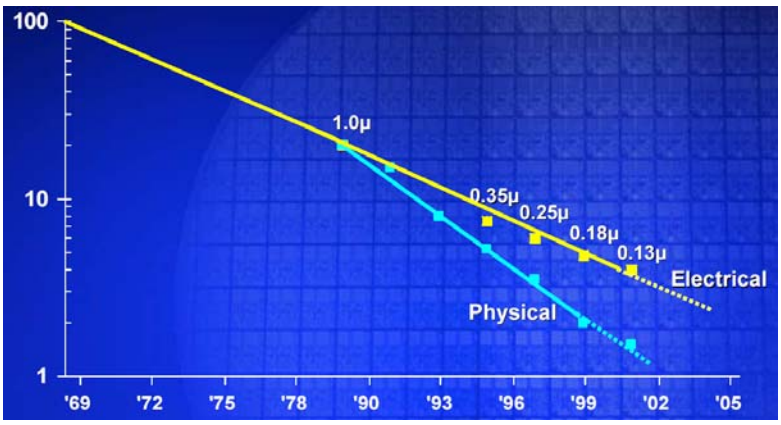
Thermal Runaway





- Destructive positive feedback mechanism
- Leakage increases exponentially with temperature
- May destroy the test socket → thermal sensors required

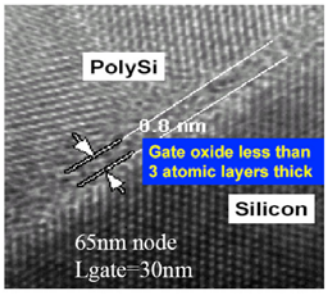
Gate Oxide Thickness

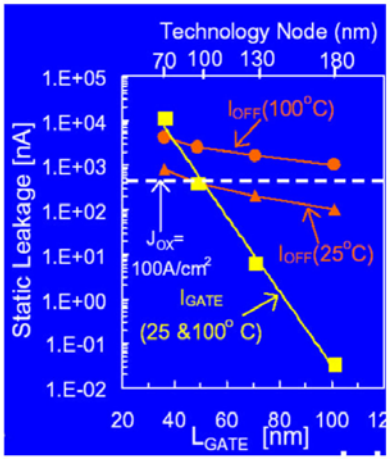


- Electrical t_{ox} > Physical t_{ox}
- Due to gate depletion and carrier quantization in the channel

Gate Tunneling Leakage

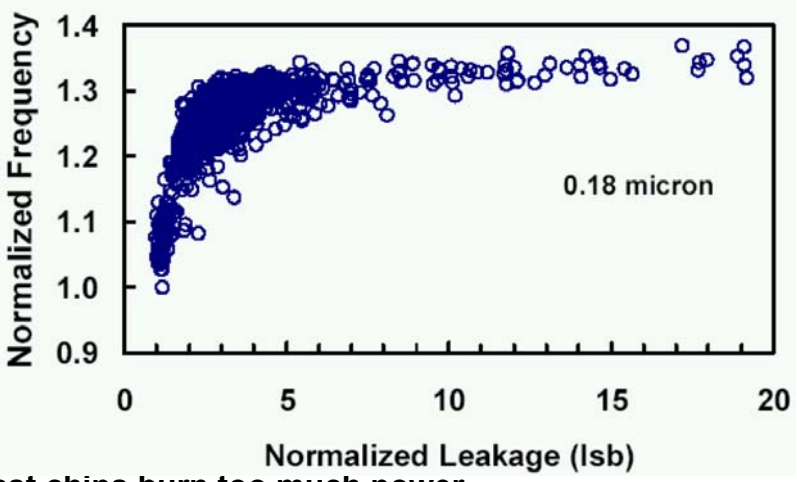
Placing a few SiO_N species uniformly in billions of devices ??





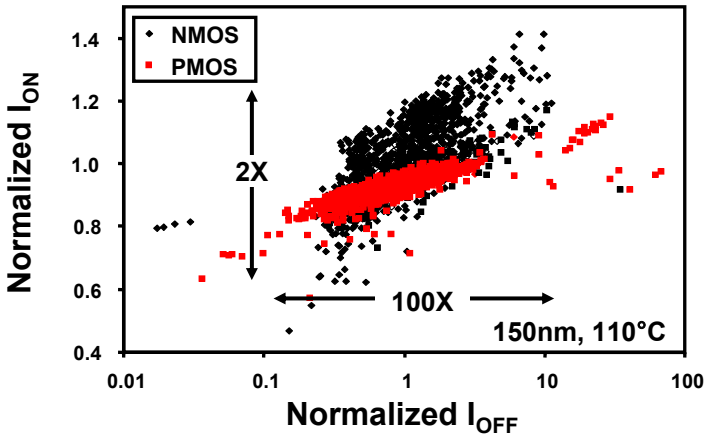
- **MOSFET no longer have infinite input resistance**
- **Impacts both power and functionality of circuits**

Process Variation in Processors



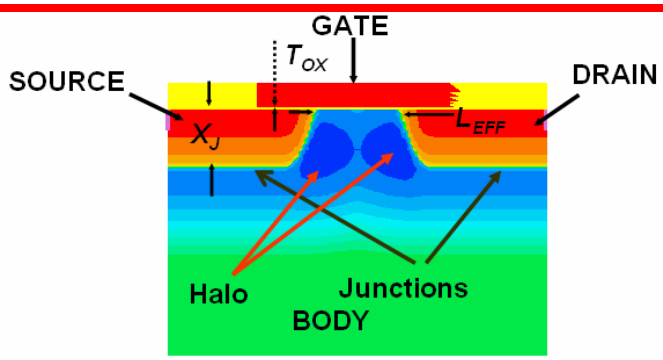
- **Fast chips burn too much power**
- **Slow chips cannot meet the frequency requirement**

Process Variation in Transistors



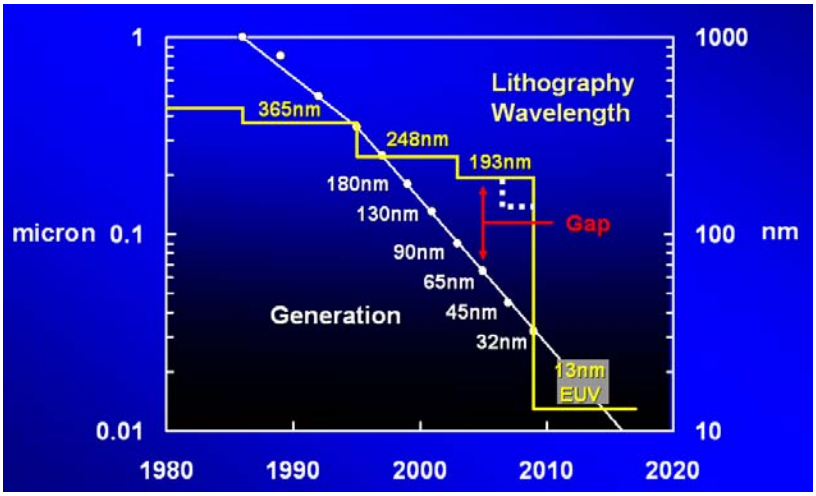
- More than 2X variation in I_{on}, 100X variation I_{off}
- Within-dies, die-to-die, lot-to-lot

Sources of Process Variation

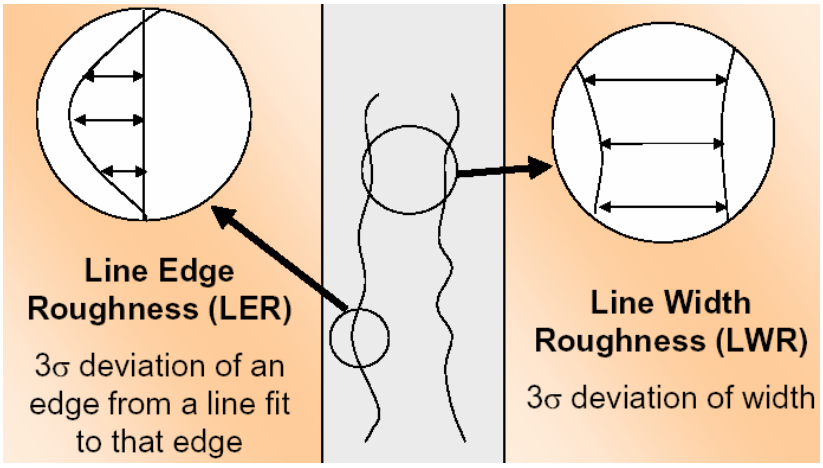


- Intrinsic parameter variation (static)
  - Channel length, random dopant fluctuation
- Environmental variation (dynamic)
  - Temperature, supply variations

Sub-wavelength Lithography

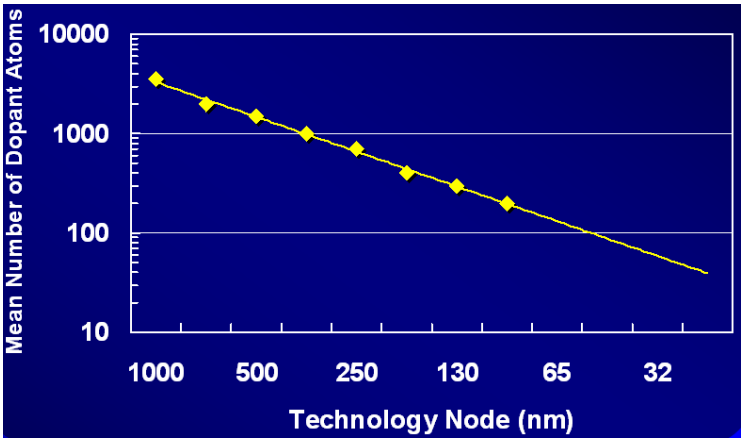


Line Edge/Width Roughness



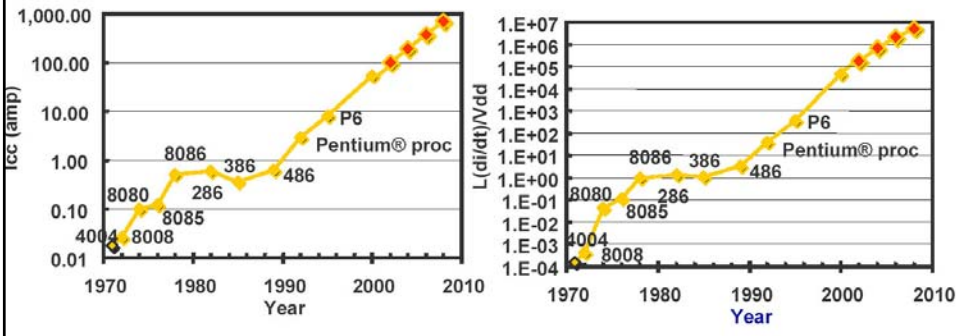
- **I_{off} and I_{dsat} impacted by LER and LWR**

Random Dopant Fluctuation



- V_t variation caused by non-uniform channel dopant distribution

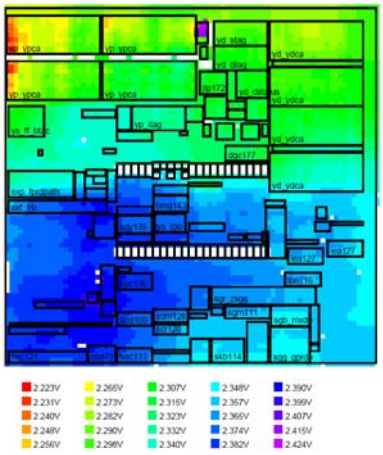
Supply Voltage Integrity



- IR noise due to large current consumption
- Ldi/dt noise due to new power reduction techniques (clock gating, power gating, body biasing) with power down mode

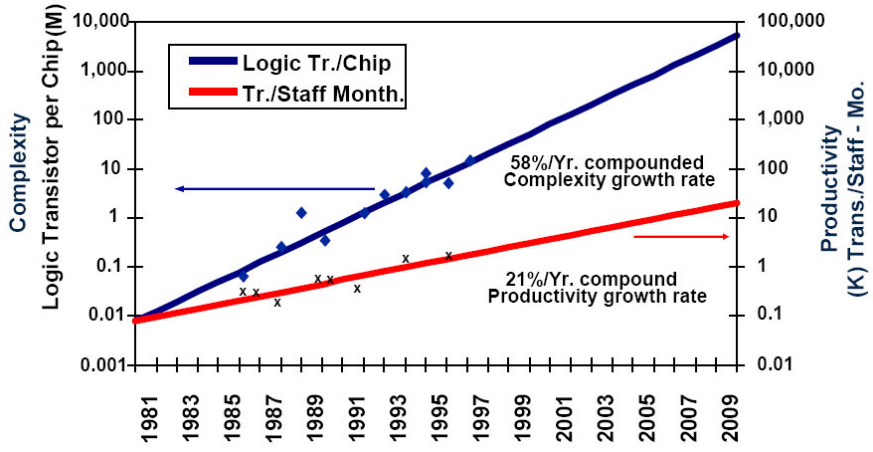
Supply Voltage Integrity

- Degrades circuit performance
- Supply voltage overshoot causes reliability issues
- Power wasted by parasitic resistance causes self-heating
- V_{dd} fluctuation should be less than 10%



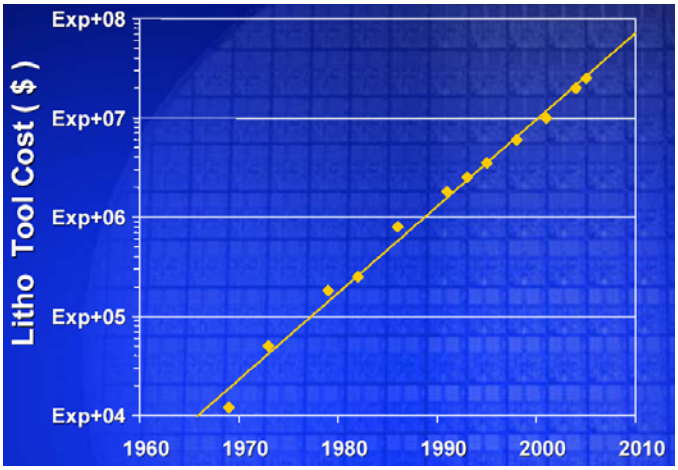
Courtesy IBM

Productivity Gap



- Design complexity surpasses manpower
- Effective CAD tools, memory dominated chips

Lithography Tool Cost



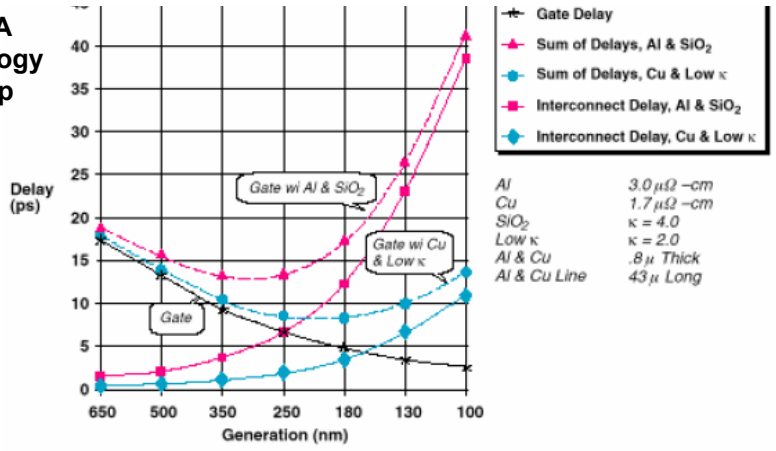
- **What will end Moore's law, economics or physics?**

Interconnect Scaling

- **Global interconnects get longer due to larger die size**
- **Wire scaling increases R, L and C**
- **Example: local vs. global interconnect delay**

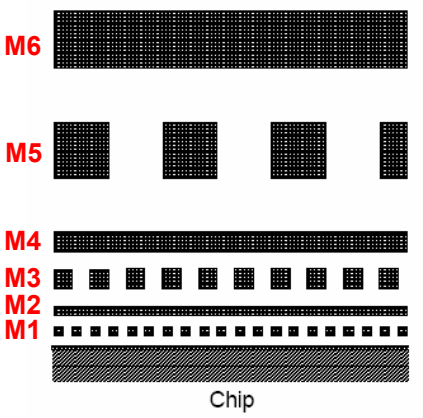
Interconnect Delay Problem

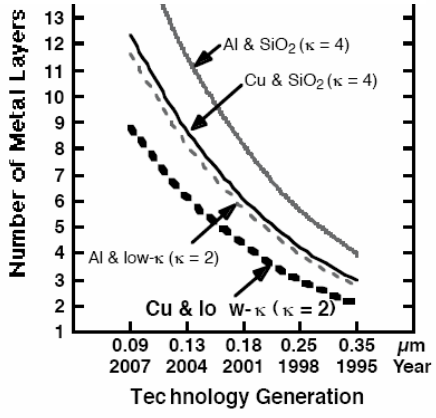
1997 SIA technology roadmap



- **Local interconnect has sped up (shorter wires)**
- **Global interconnect has slowed down (RC doesn't scale)**

Interconnect Metal Layers

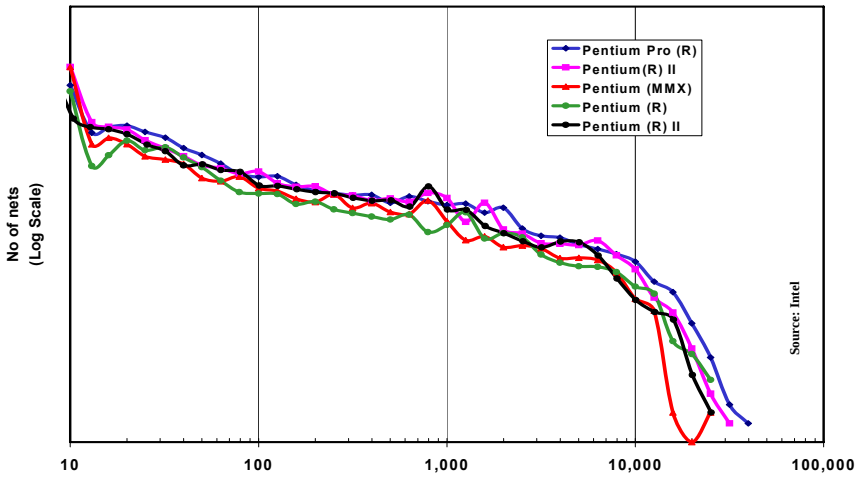




- **Local wires have high density to accommodate the increasing number of devices**
- **Global wires have low RC (tall, wide, thick, scarce wires)**

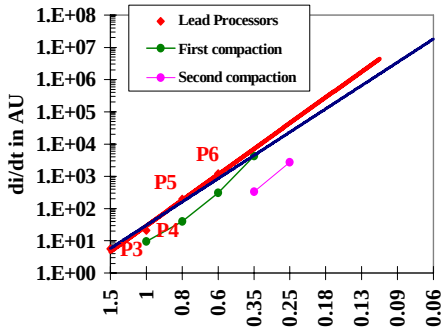
Interconnect distribution scaling trends

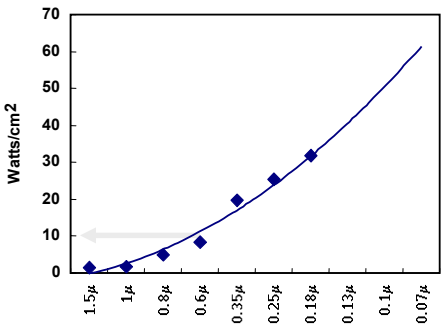
- RC/μm scaling trend is only one side of the story...



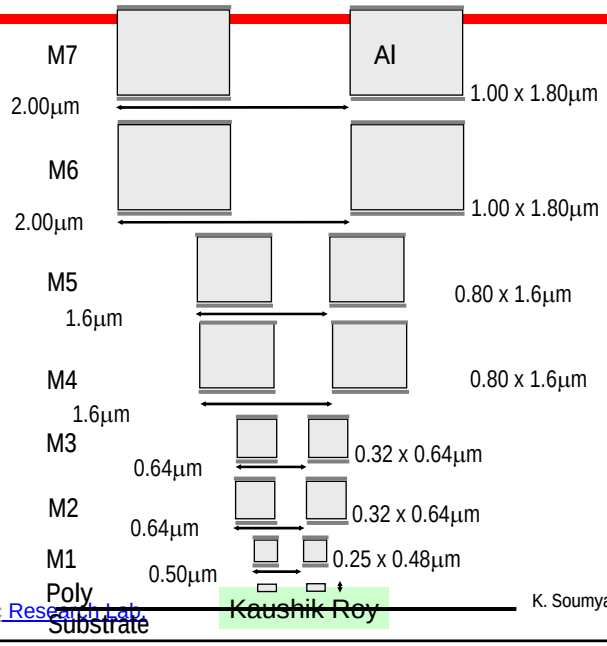
Power Delivery & Distribution Challenges

- **High-end microprocessors approaching > 10 GHz**
  - **How to deliver and distribute ~100A at < 1V for < $20!**
  - **On-die power density >>> hot-plate power density**
    - **crossover happened back in 0.6μm technology!**
  - **di/dt noise only worsening with scaling: drivers are one of the sources.**

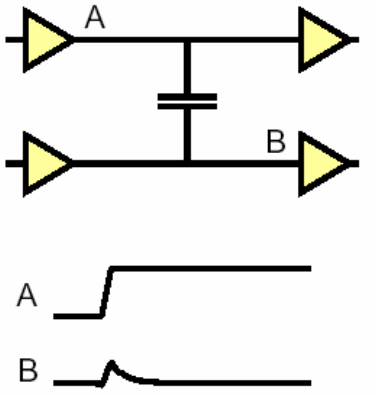




Example multi-layer system

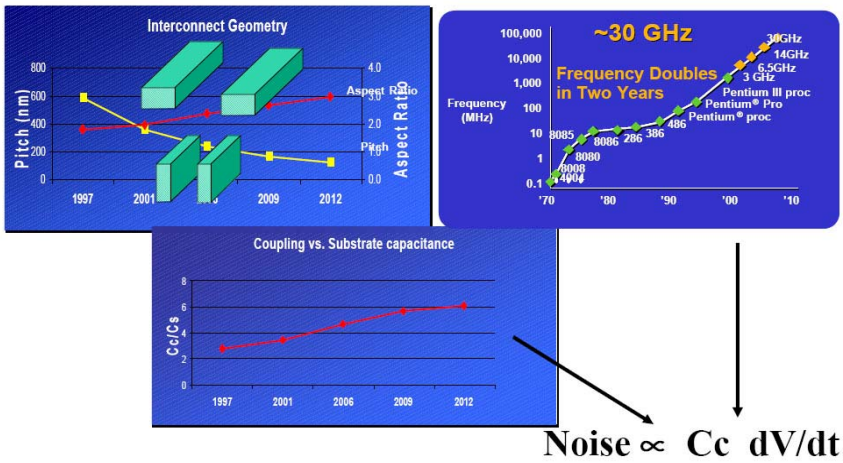


Cross Talk Noise



- **As wires are brought closer with scaling, capacitive coupling becomes significant**
- **Adjacent wires on same layer have stronger coupling**

Cross Talk Noise

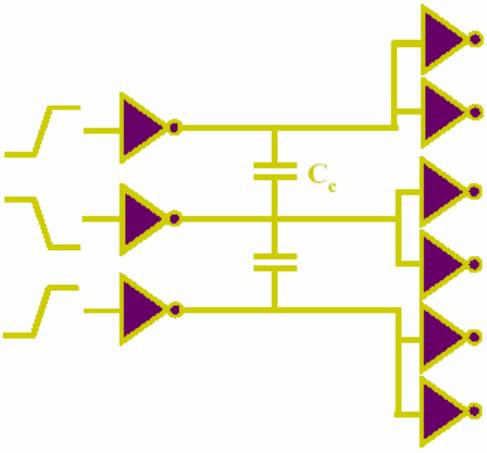


$$\text{Noise} \propto Cc\ dV/dt$$

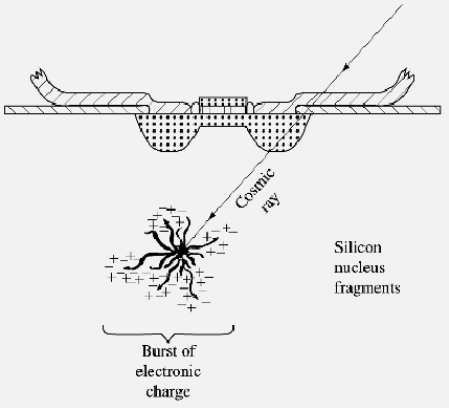
- **Multiple aggressors multiple victims possible**
- **Cross talk noise can cause logic faults in dynamic circuits**

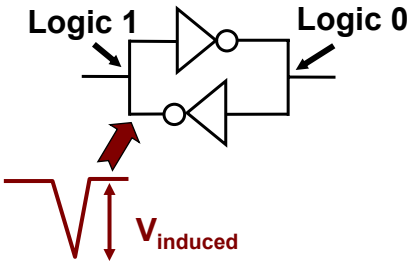
Cross Talk and Delay

- **Capacitive cross talk can affect delay**
- **If aggressor(s) switch in opposite direction, effective coupling capacitance is doubled**
- **On the other hand, if aggressor(s) switch in the same direction, Cc is eliminated**
- **Significant difference in RC delay depending on adjacent switching activity**



Soft Error In Storage Nodes

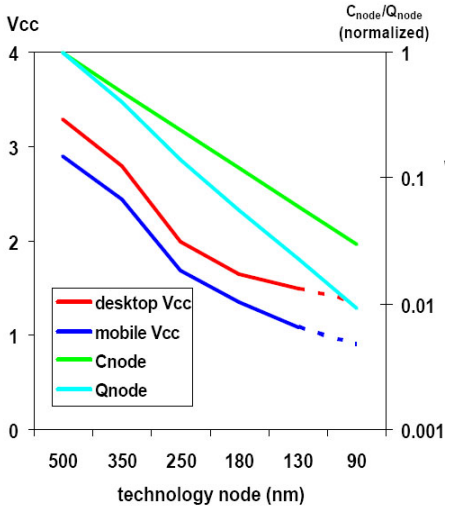




- **Soft errors are caused by**
  - » **Alpha particles from package materials**
  - » **Cosmic rays from outer space**

Soft Error In Storage Nodes

- **Error correction code**
- **Shielding**
- **SOI**
- **Radiation-hardened cell**



More Roadblocks

- Memory stability
- Long term reliability
- Mixed signal design issues
- Mask cost
- Testing multi-GHz processors
- Skeptics: Do we need a faster computer?
- …

- Eventually, it all boils down to economics

Summary

- Digital IC Business is Unique
  - Things Get Better Every Few Years
  - Companies Have to Stay on Moore's Law Curve to Survive
- Benefits of Transistor Scaling
  - Higher Frequencies of Operation
  - Massive Functional Units, Increasing On-Die Memory
  - Cost/MIPS Going Down
- Downside of Transistor Scaling
  - Power (Dynamic and Static)
  - Process Variation
  - Design/Manufacturing Cost
  - ….