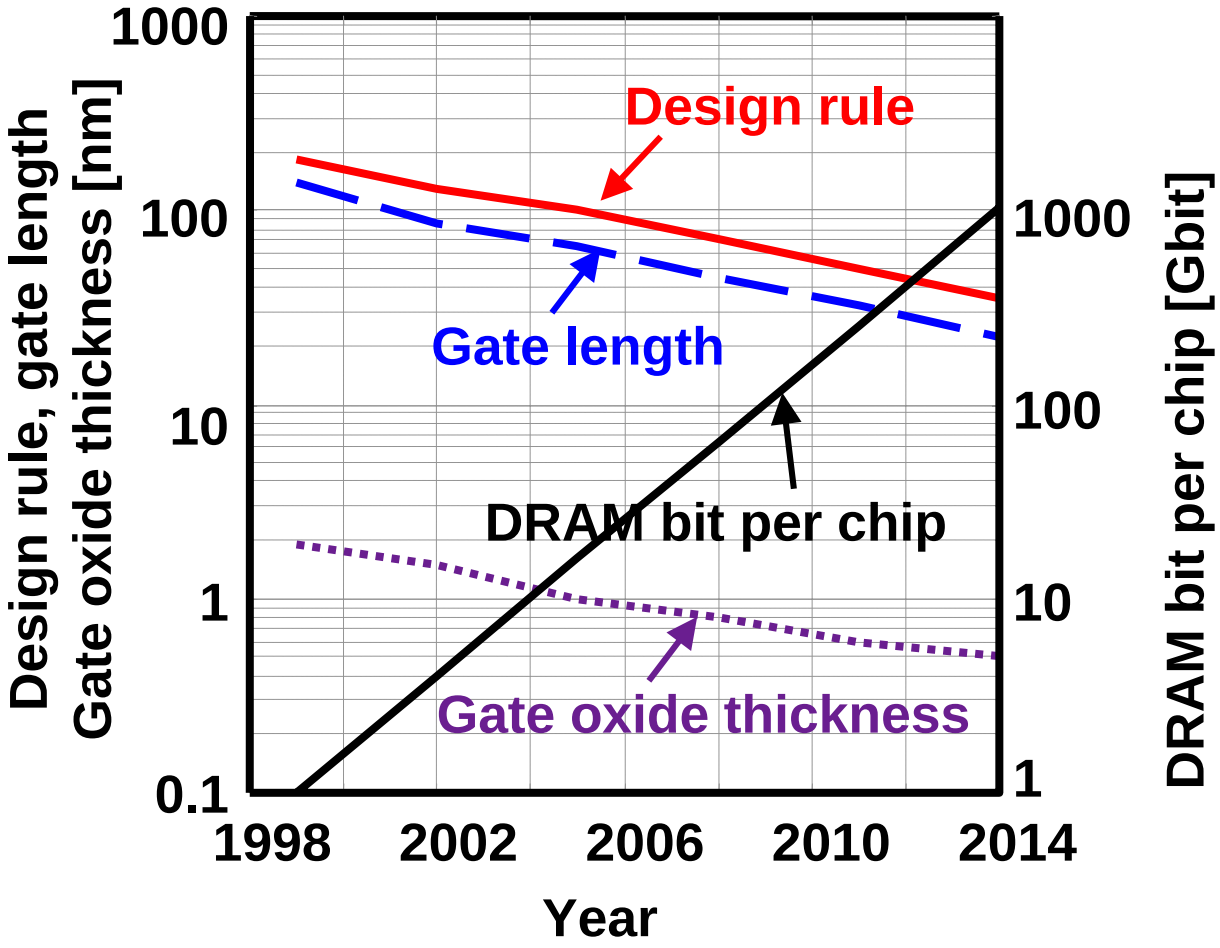
Design closure forum, Feb. 1st, 2001

VLSI design challenges in the forthcoming decade

- What may hinder design closure -

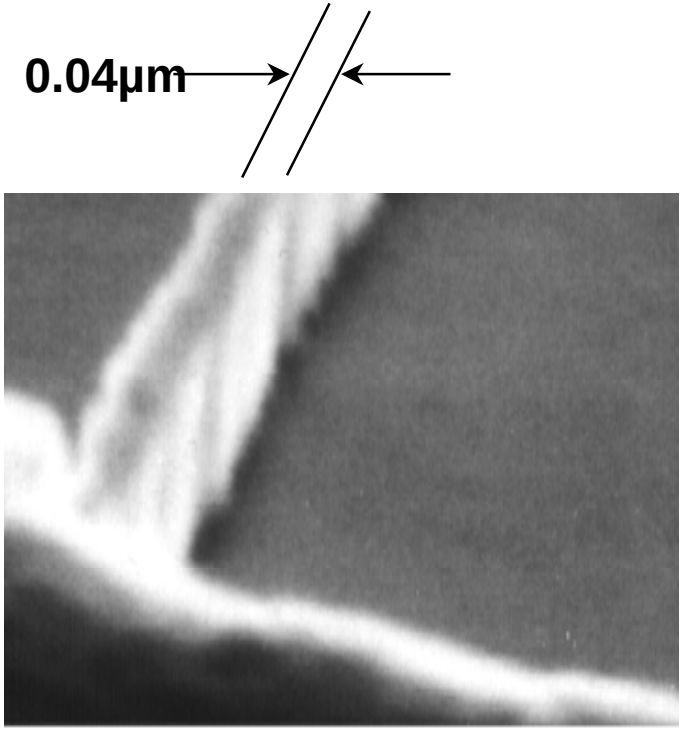
Prof. Takayasu Sakurai
Center for Collaborative Research, and
Institute of Industrial Science,
University of Tokyo
E-mail:tsakurai@iis.u-tokyo.ac.jp

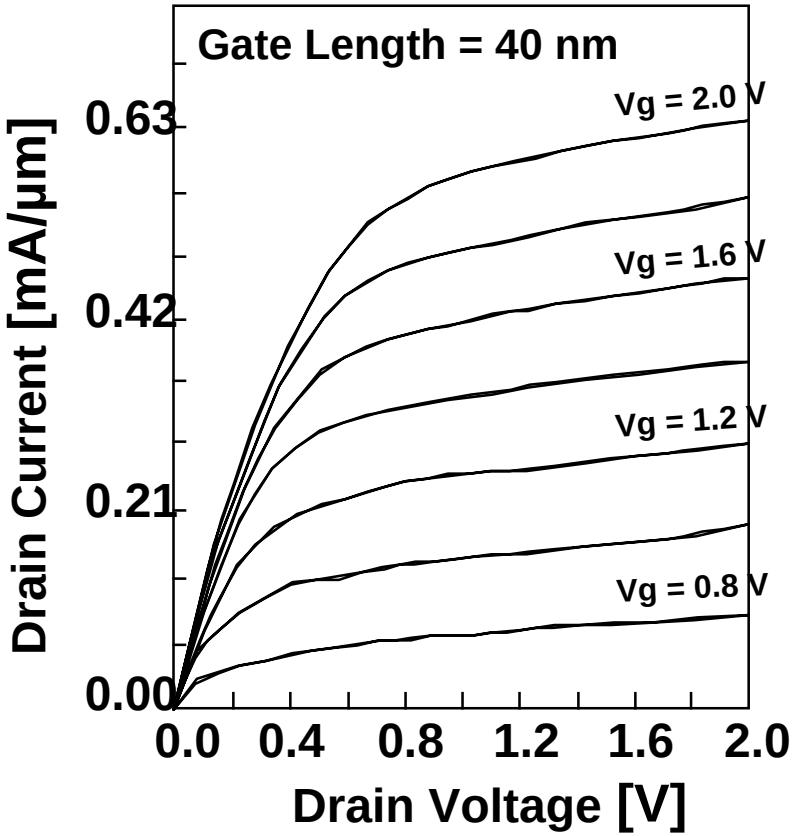
Technology trend



Limit of Miniturization

0.04μm MOSFET



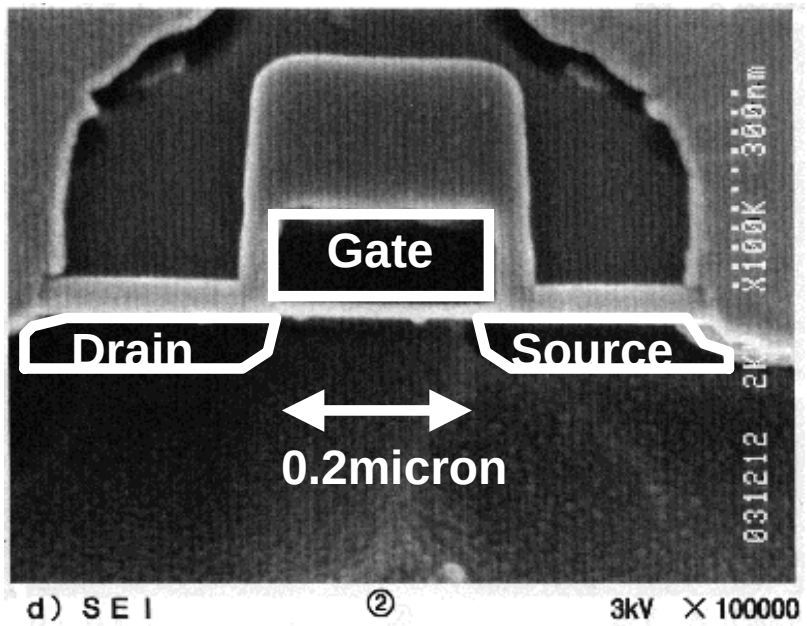


Conventional I-V curve at 0.04μm (Even down to 0.014μm)

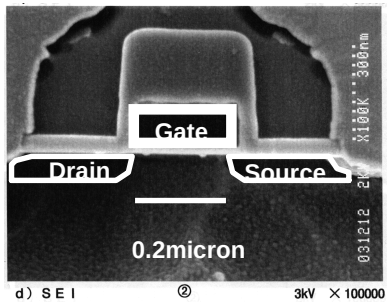
M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, and H. Iwai, "Sub-50nm gate Length N-MOSFETs with 10 nm Phosphorus Source and Drain Junctions", IEDM Technical Digest, pp. 119 - 122, 1993.

H. Kawaura, T. Sakamoto, Y. Ochiai, J. Fujita, and T. Baba, "Fabrication and Characterization of 14-nm-Gate-Length EJ-MOSFETs", Extended Abstracts of SSDM, pp.572-573, 1997.

Scaling Law



Size 1/2



Favorable effects

Size	x1/2
Voltage	x1/2
Electric Field	x1
Speed	x3
Cost	x1/4

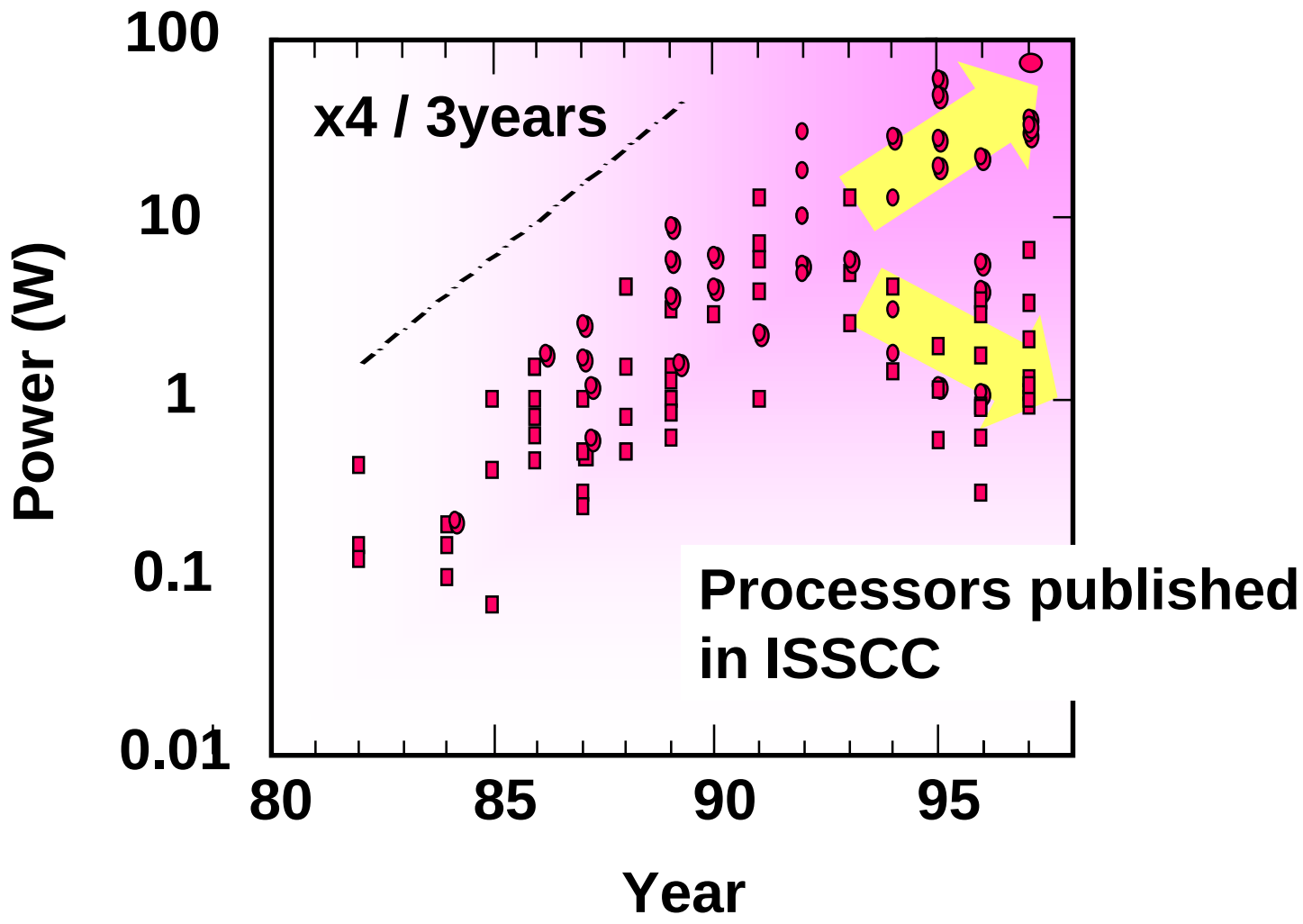
Unfavorable effects

Power density	x1.6
RC delay/Tr. delay	x3.2
Current density	x1.6
Voltage noise	x3.2
Design complexity	x4

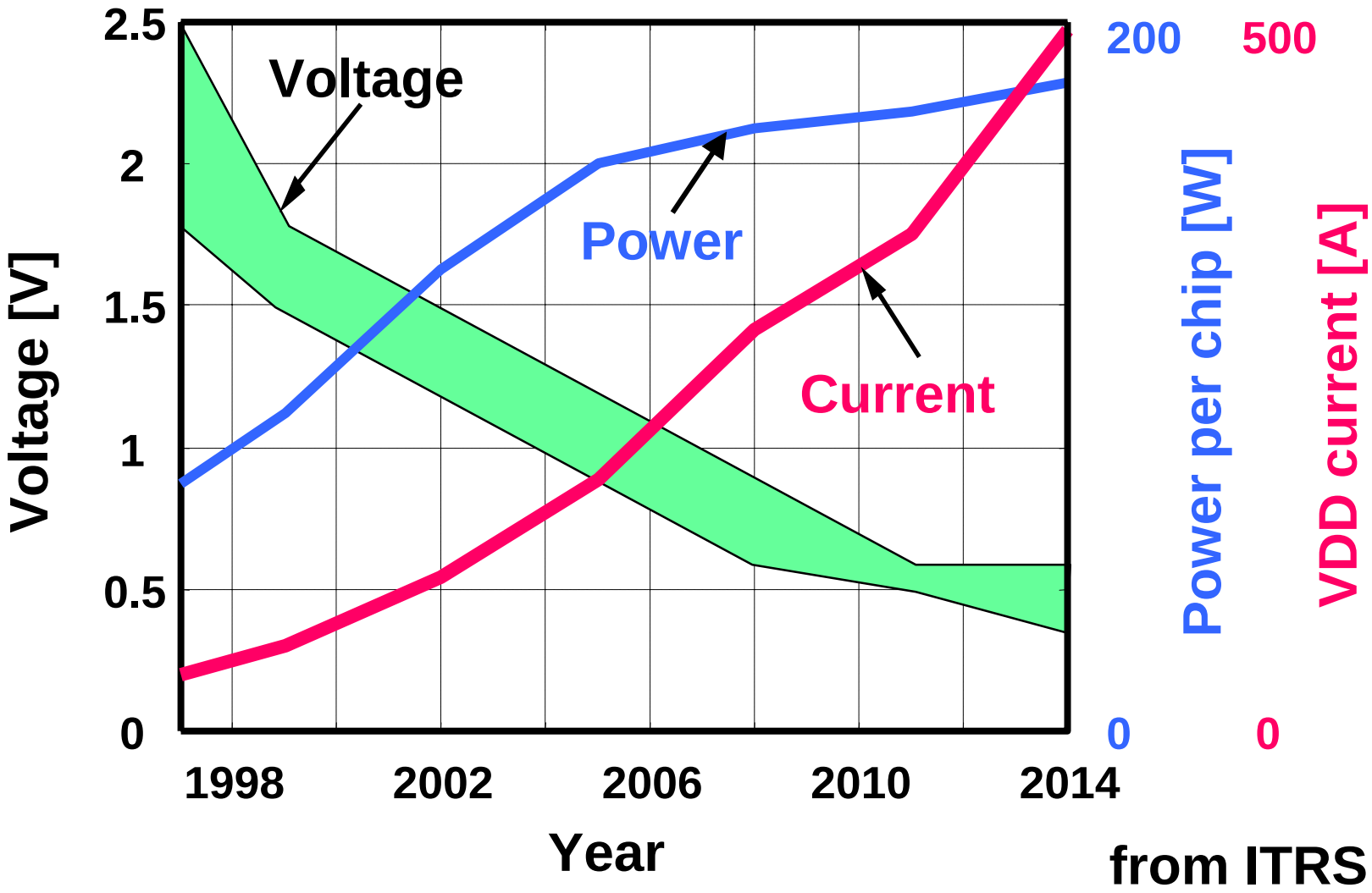
Three crises in VLSI designs

- Power crisis
- Interconnection crisis
- Complexity crisis

Ever Increasing VLSI Power



V_{DD}, power and current trend

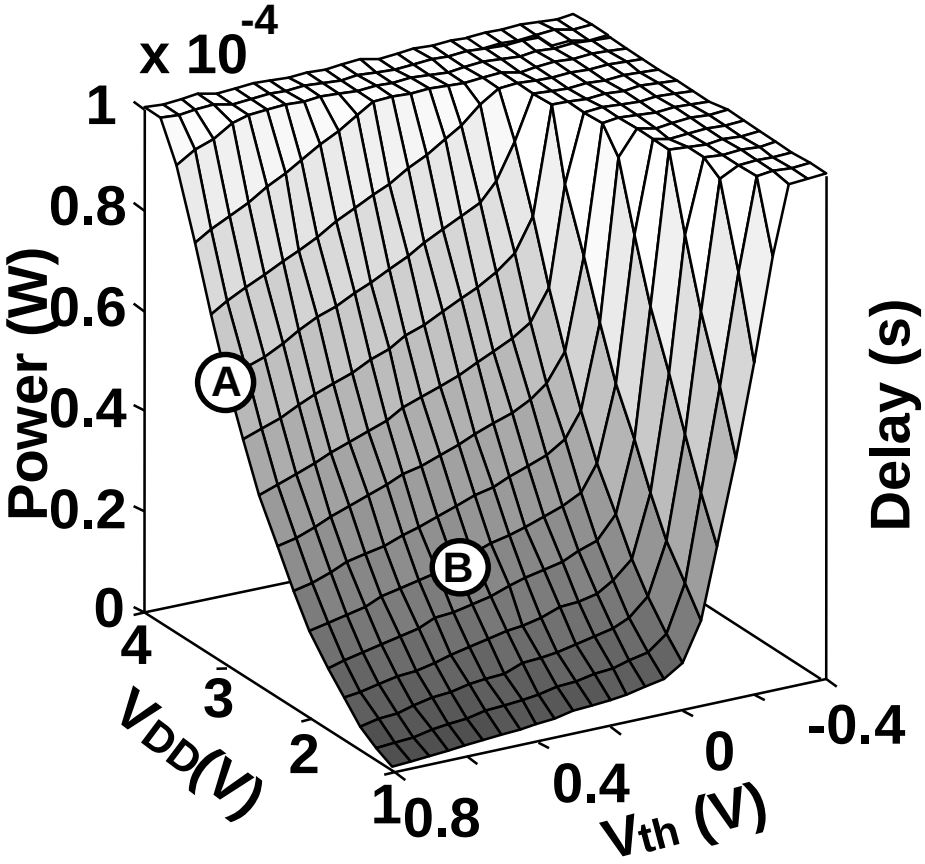


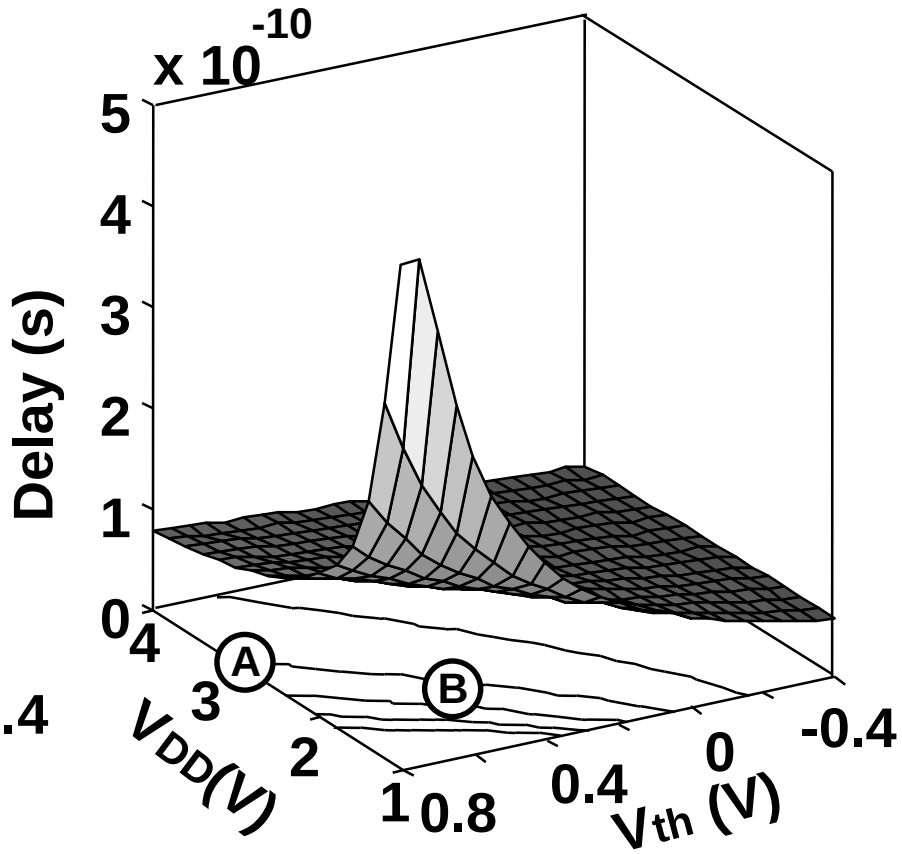
from ITRS

Power & Delay Dependence on V_{DD} & V_{TH}

Power : $P = p_t \cdot f_{CLK} \cdot C_L \cdot V_{DD}^2 + I_0 \cdot 10^{-\frac{V_{th}}{S}} \cdot V_{DD}$

Delay $= \frac{k \bullet Q}{I} = \frac{k \cdot C_L \cdot V_{DD}}{(V_{DD} - V_{th})^{\alpha}}$ $(\alpha=1.3)$





Controlling V_{DD} and V_{TH} for low power

Low power → Low V_{DD} → Low speed → Low V_{TH} → High leakage → V_{DD}-V_{TH} control

	Active	Stand-by
Multiple V_{TH}	Dual-V_{TH}	MTCMOS
Variable V_{TH}	V_{TH} hopping	VTCMOS
Multiple V_{DD}	Dual-V_{DD}	Boosted gate MOS
Variable V_{DD}	V_{DD} hopping	

Software-hardware cooperation

Technology-circuit cooperation

*) MTCMOS: Multi-Threshold CMOS

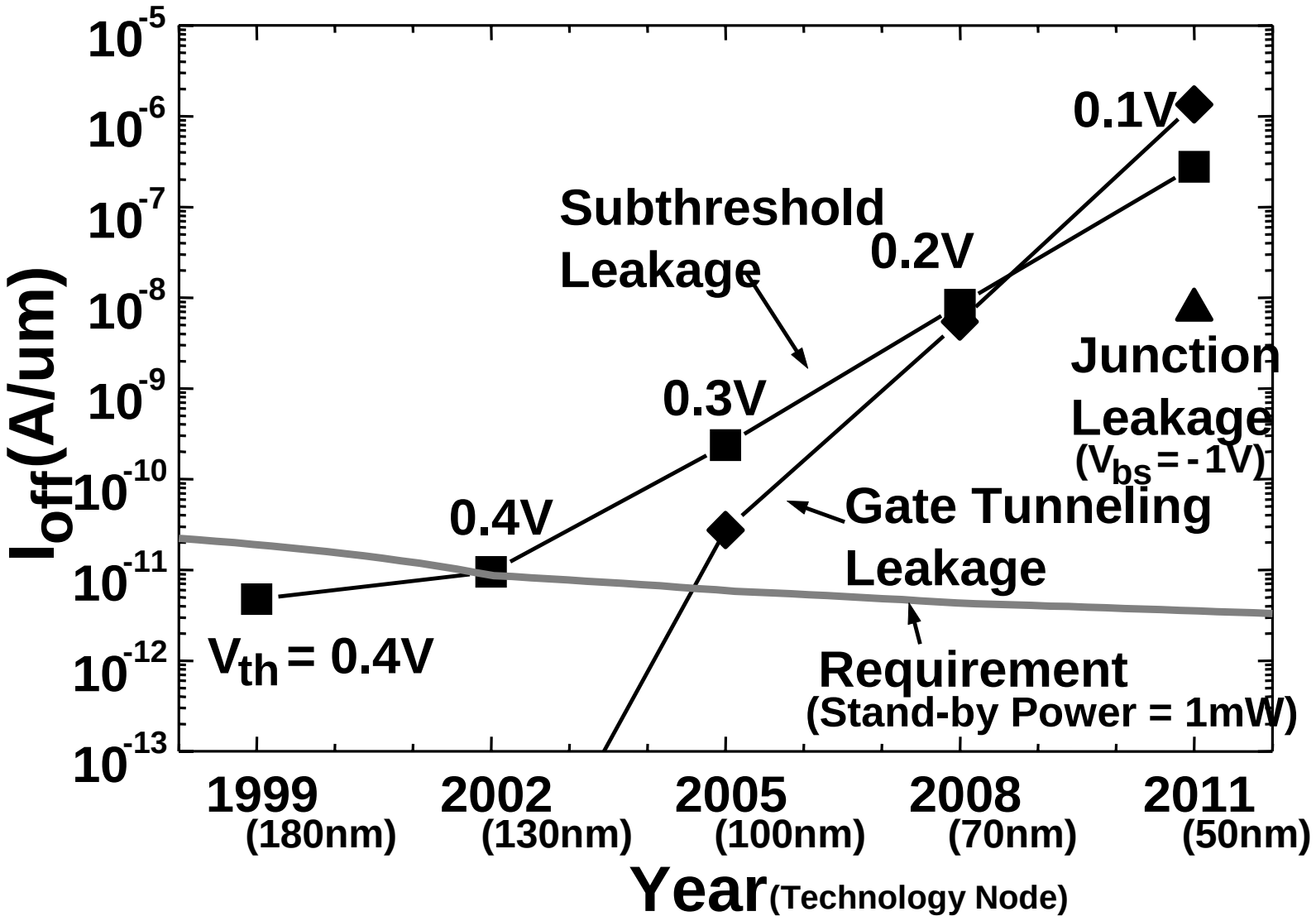
*) VTCMOS: Variable Threshold CMOS

- Multiple : spatial assignment
- Variable : temporal assignment

Controlling V_{DD} and V_{TH} for low power

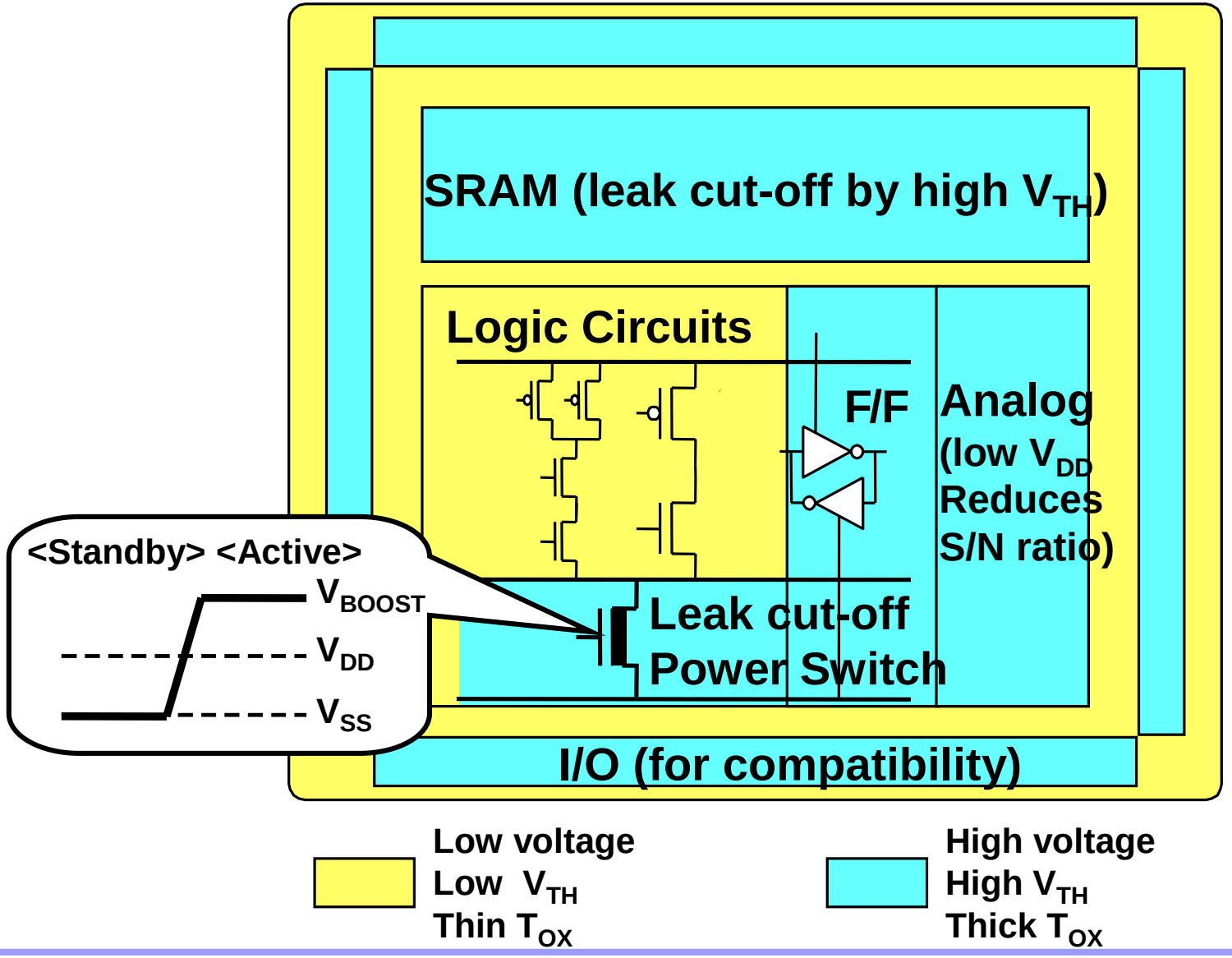
	Active	Stand-by
Multiple V_{TH}	Dual-V_{TH}	MTCMOS
Variable V_{TH}	VTCMOS	VTCMOS
Multiple V_{DD}	Dual-V_{DD}	Boosted gate MOS
Variable V_{DD}	V_{DD} hopping	

Transistors go leaky

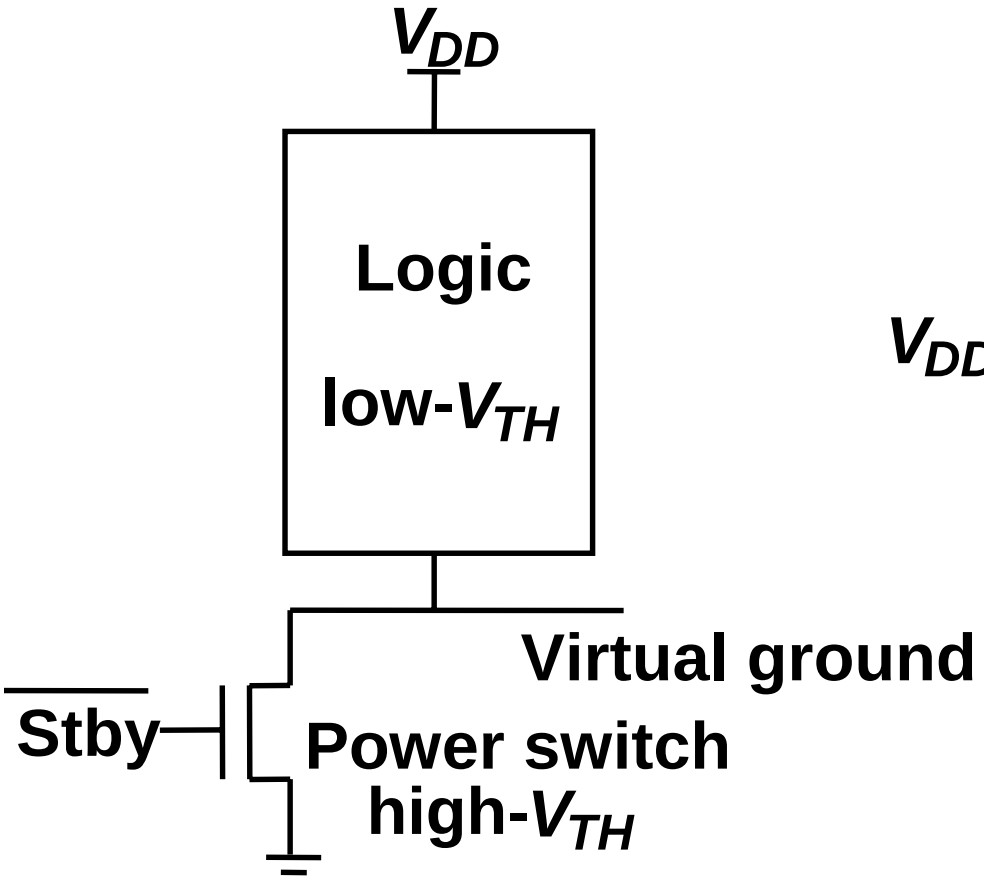


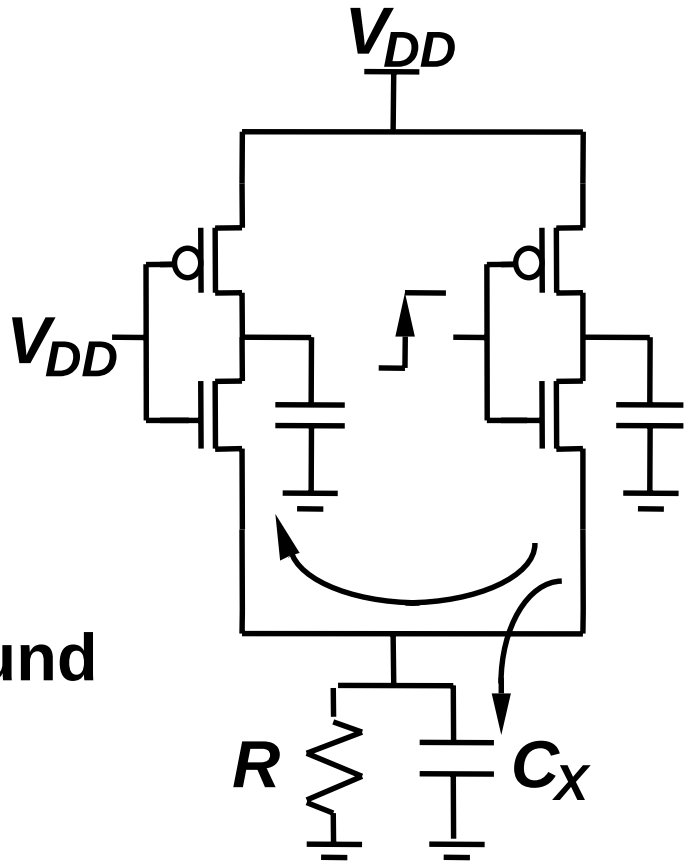
T.Sakurai

GSI's in deep-submicron era



Power switch gate width in BGMOS





Kao, DAC'97, pp.409-414.

Degrade circuit speed unpredictably

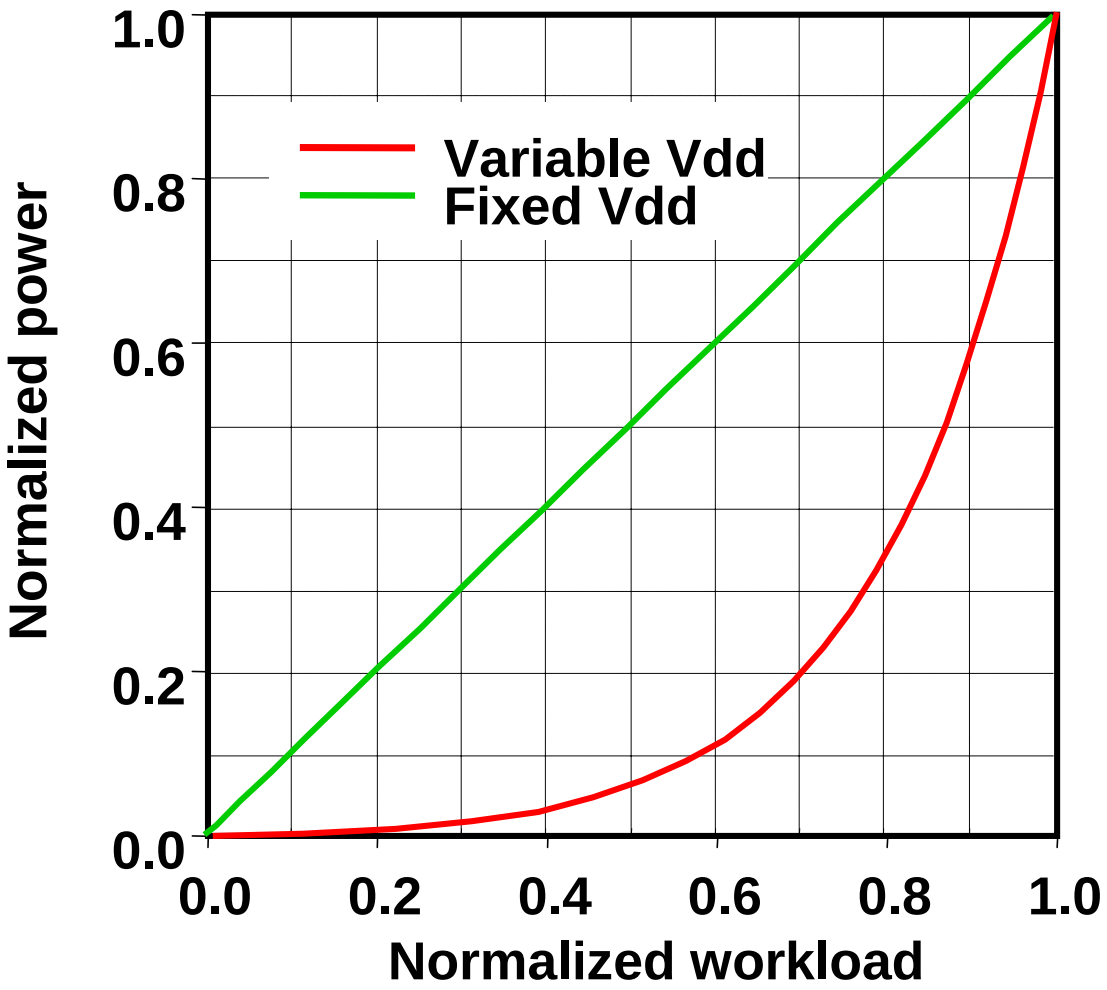
Controlling V_{DD} and V_{TH} for low power

	Active	Stand-by
Multiple V_{TH}	Dual-V_{TH}	MTCMOS
Variable V_{TH}	V_{TH} hopping	VTCMOS
Multiple V_{DD}	Dual-V_{DD}	Boosted gate MOS
Variable V_{DD}	V_{DD} hopping	

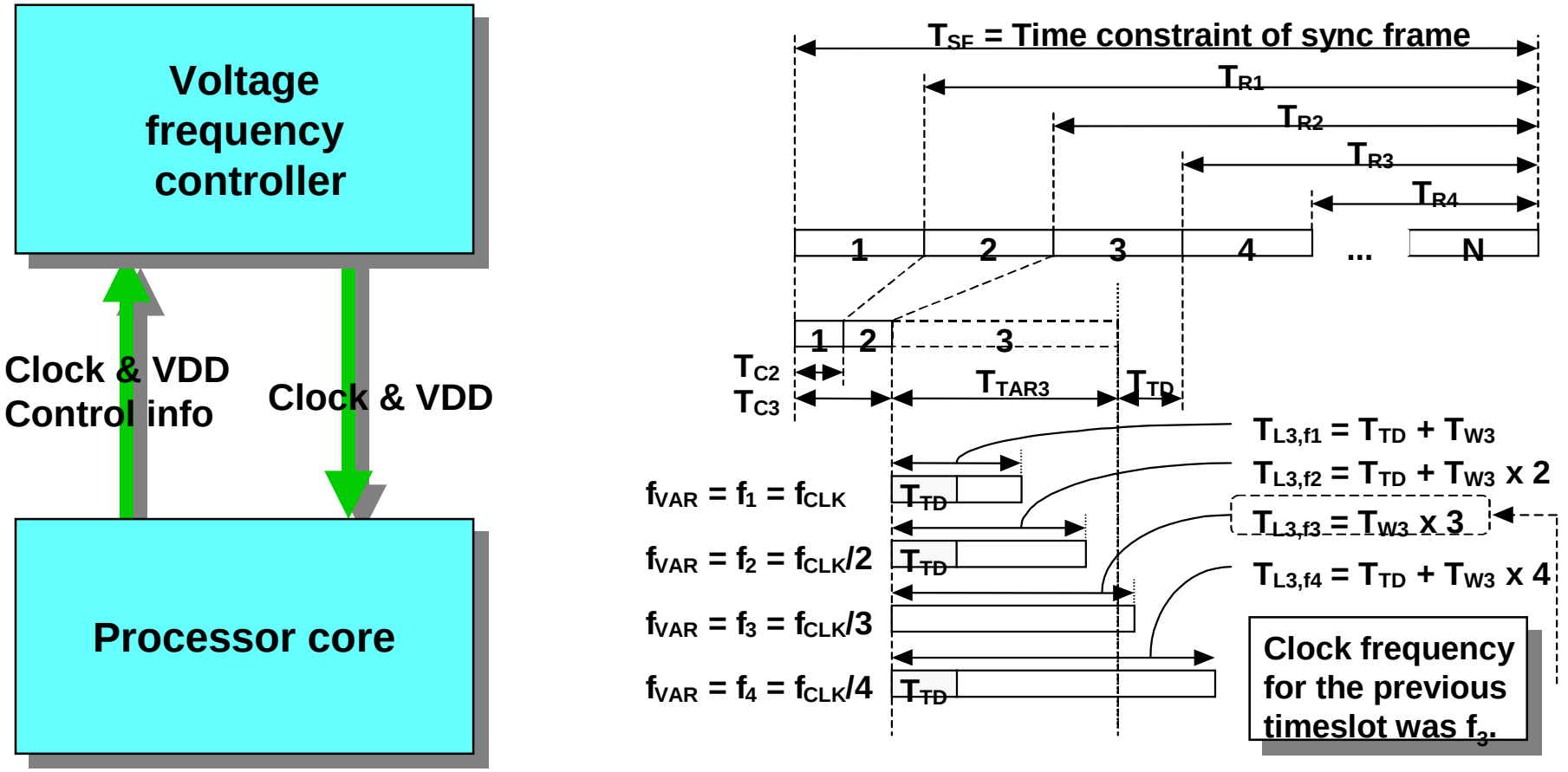
If you don't need to hussle, V_{DD} should be as low as possible

Energy consumption is proportional to the square of V_{DD}.

V_{DD} should be lowered to the minimum level which ensures the real-time operation.



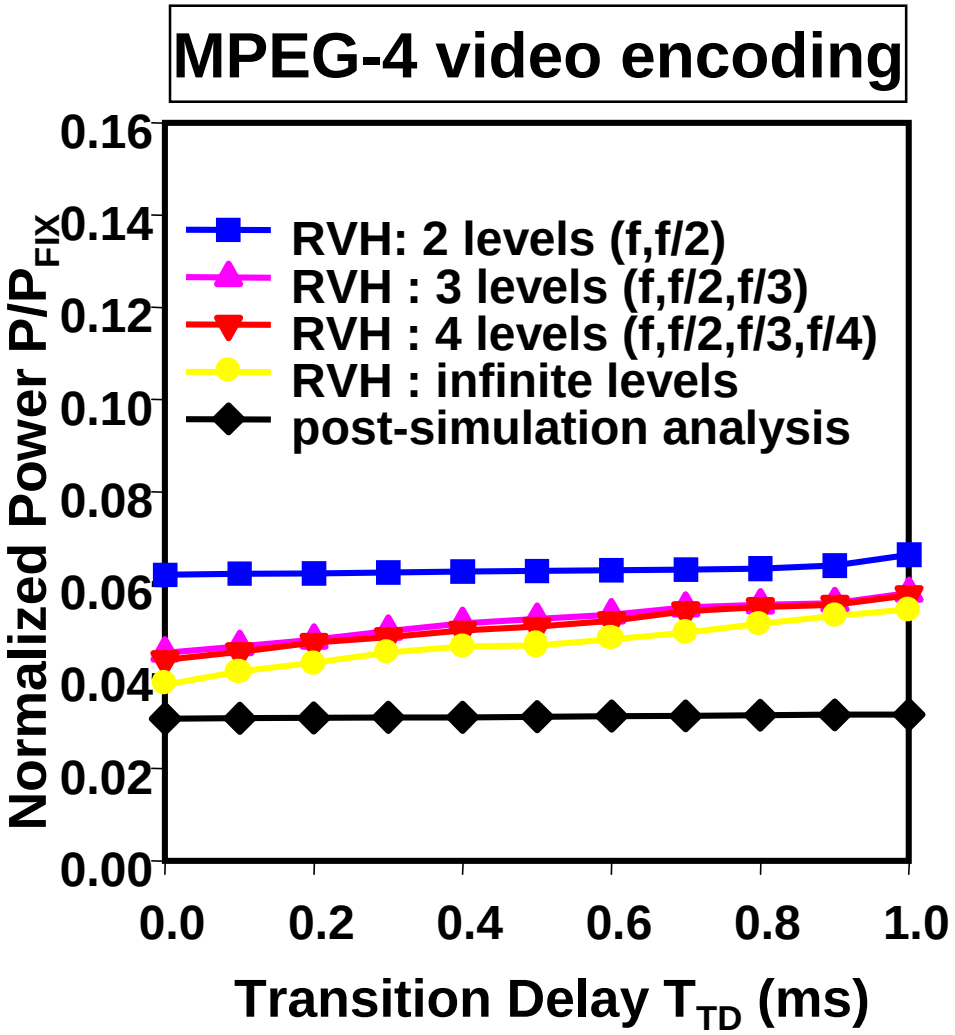
Application slicing and software feedback loop in Voltage Hopping

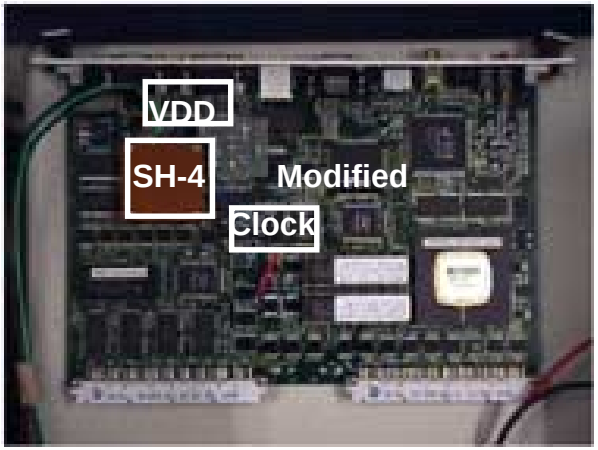


S.Lee and T.Sakurai, “Run-time Power Control Scheme Using Software Feedback Loop for Low-Power Real-time Applications,”ASPDAC'00, A5.2, pp.381~pp.386, Jan. 2000.

S.Lee and T.Sakurai, “Run-time Voltage Hopping for Low-power Real-time Systems,” DAC'00, June 2000.

Run-time Voltage Hopping reduces power to less than 1/10

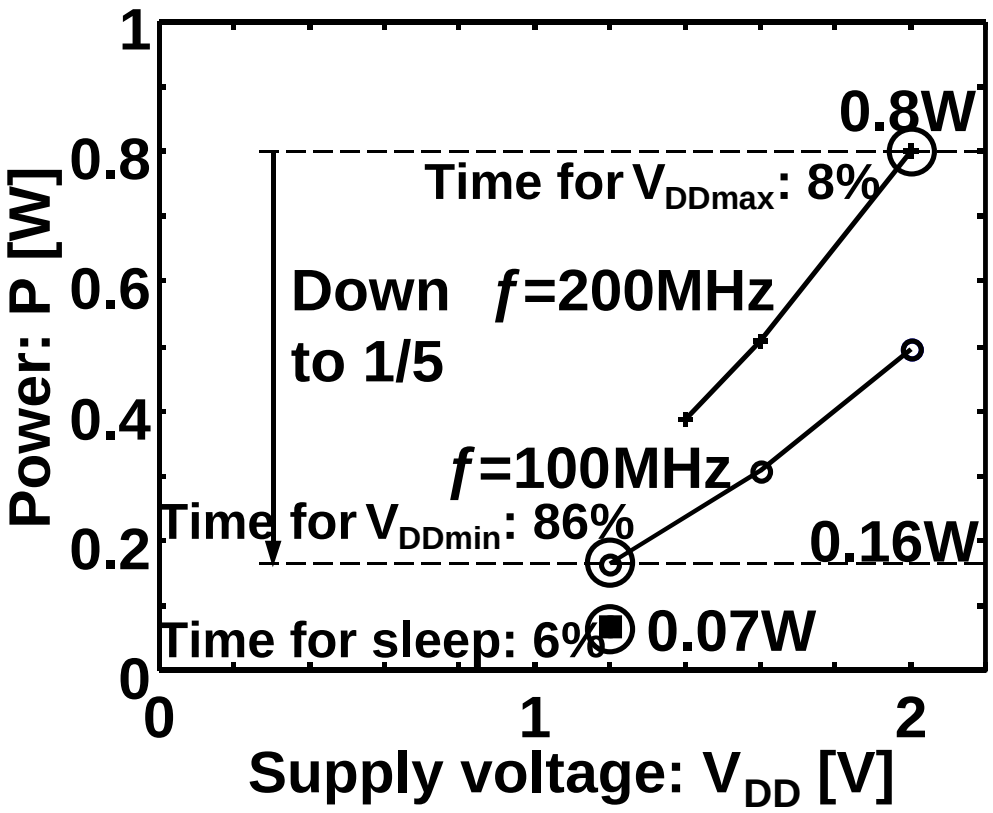




T.Sakurai

Measured power characteristics

Total power = 0.8 x 0.08 + 0.16 x 0.86 + 0.07 x 0.06 = 0.2W



VDD hopping can cut down power consumption to 1/4

Summary – low-power

- For reducing standby power, insert a power switch in series to logic circuit (BGMOS).
  → Choice of power switch gate width

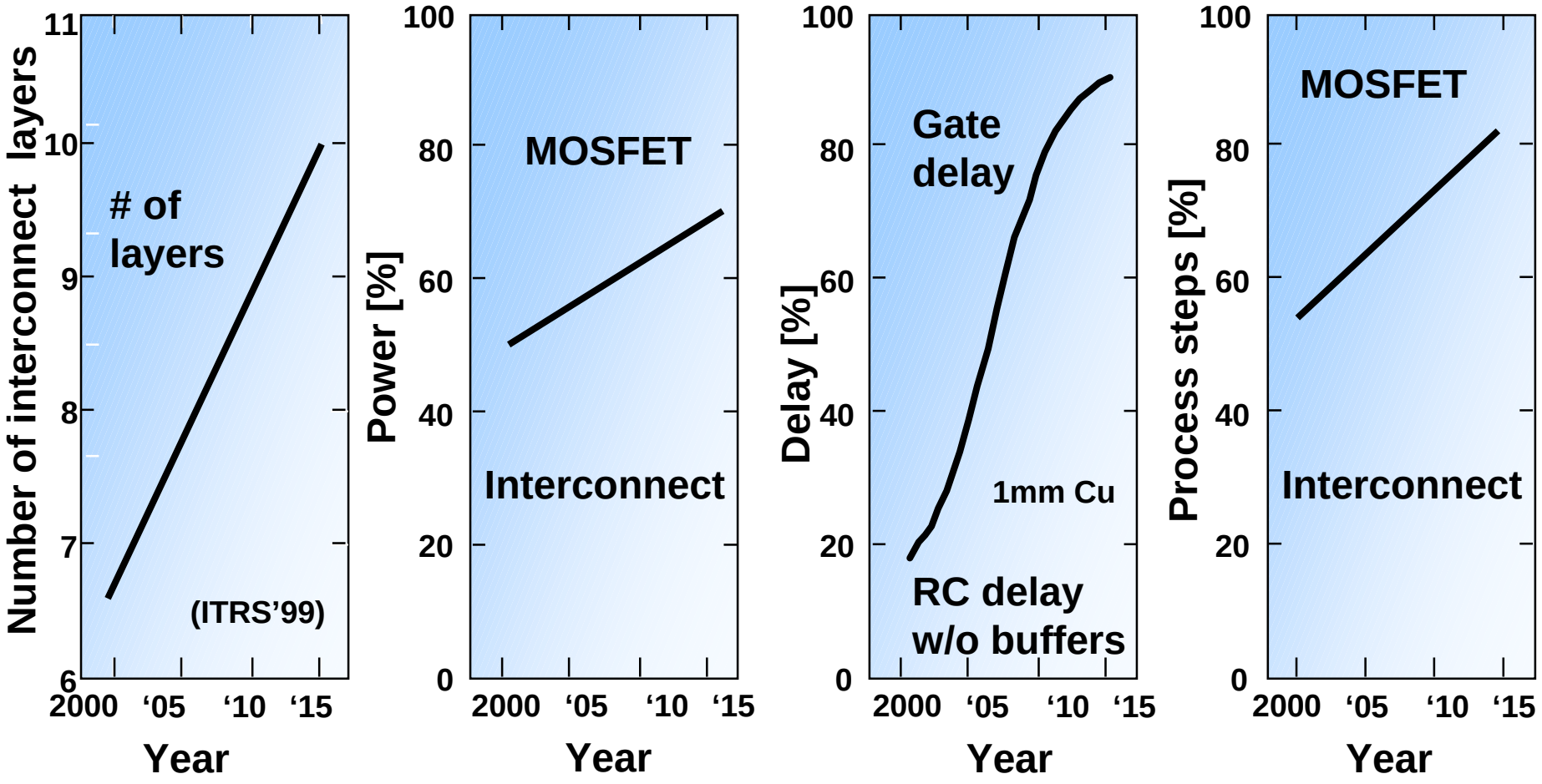
- For reducing active power, dual-VTH scheme and software control of VDD and VTH are promising.
  → Tools to support system-level low-power design with S/H co-design capability

- Future giga-scale integration will use multiple VDD, VTH and Tox.
  → Tools to support new tech.

Three crises in VLSI designs

- Power crisis
- Interconnection crisis
- Complexity crisis

Interconnect determines cost & perf.



T.Sakurai

DSM interconnect design issues

Larger current

- IR drop (static and dynamic)
- Reliability (electro-migration)

Smaller geometry / Denser pattern

- RC delay
- Signal Integrity
- Crosstalk noise
- Delay fluctuation

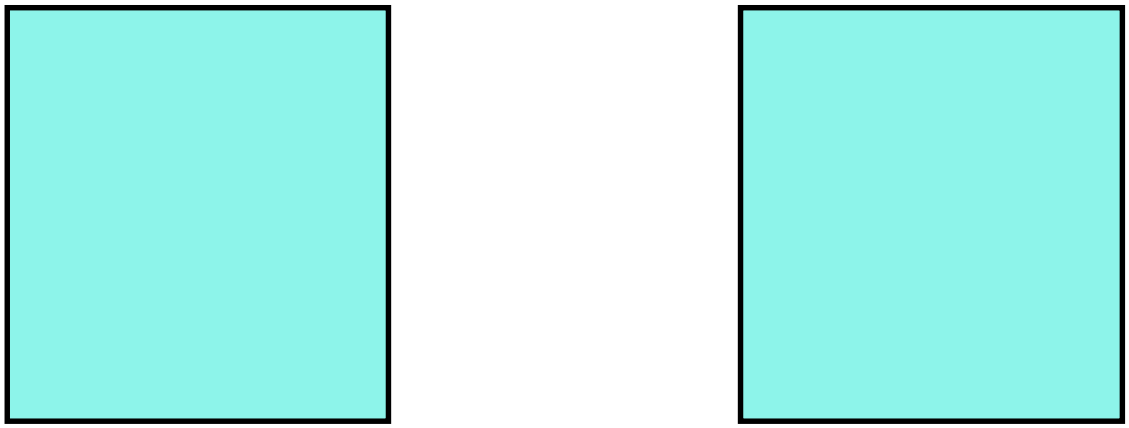
Higher speed

- Inductance
- EMI

Interconnect Cross-Section and Noise

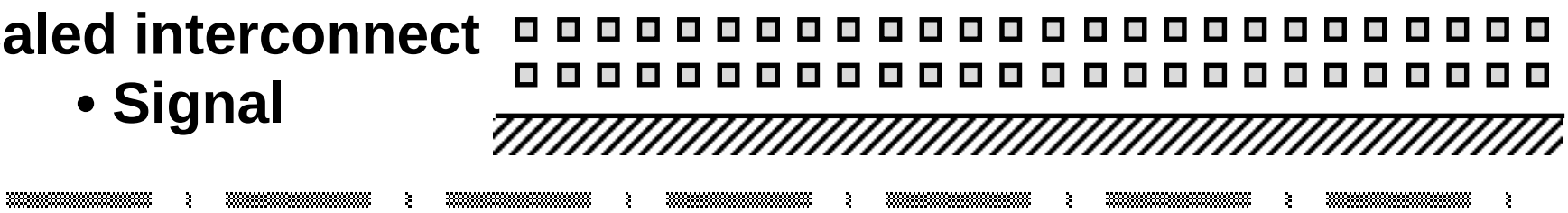
Unscaled / anti-scaled

- **Clock**
- **Long bus**
- **Power supply**



Scaled interconnect

- **Signal**

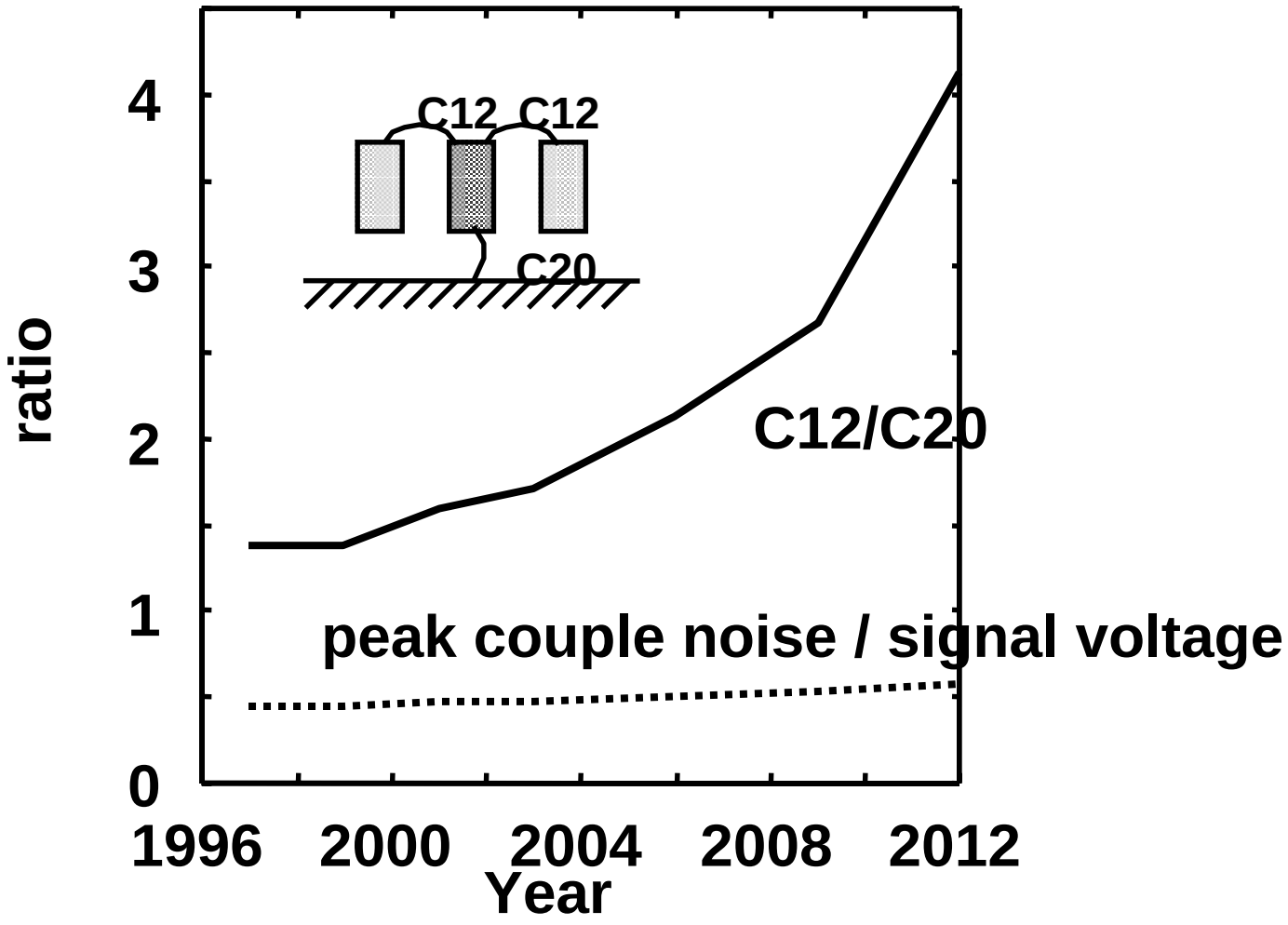


1V 20W →20A current

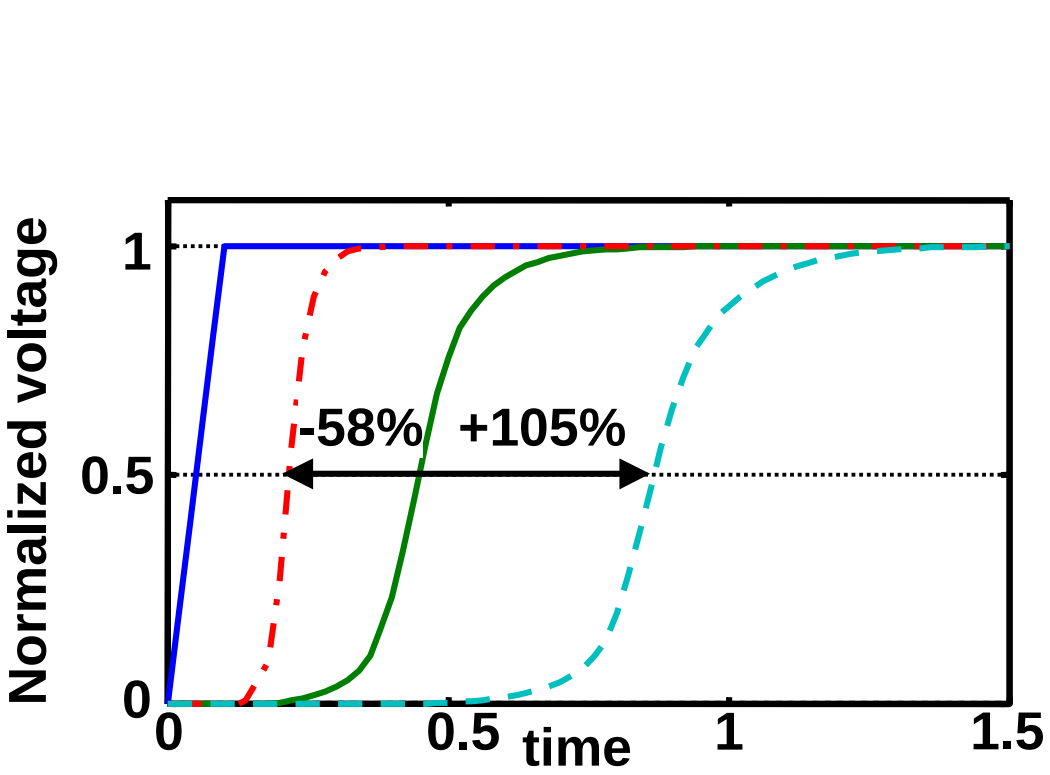
5% noise →0.05V noise →~0.02V / 20A→~10μm thick Cu

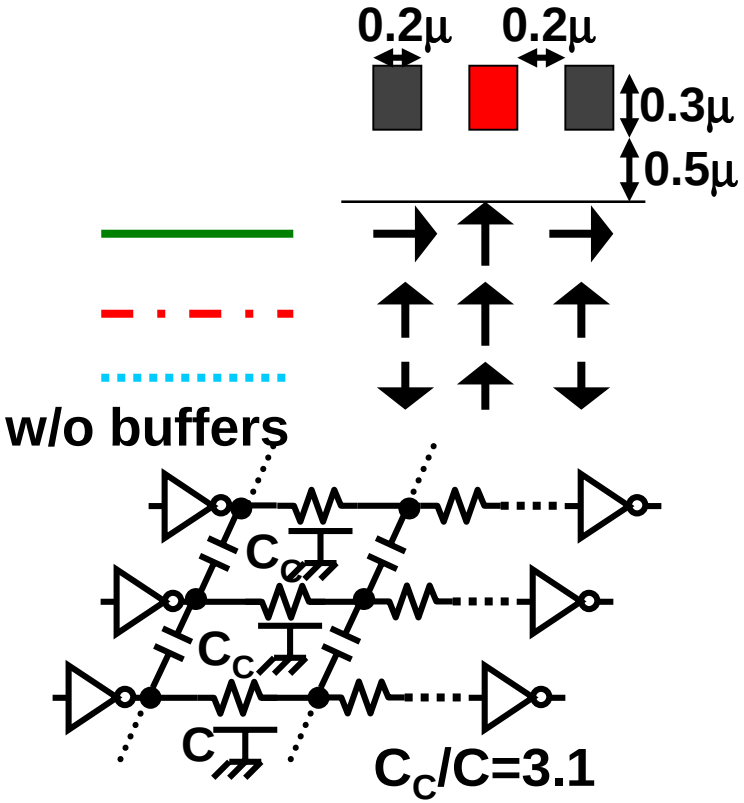
Thick layer interconnect, area pad, package are co-designed.

Capacitive Coupling Noise



Coupling among Interconnections



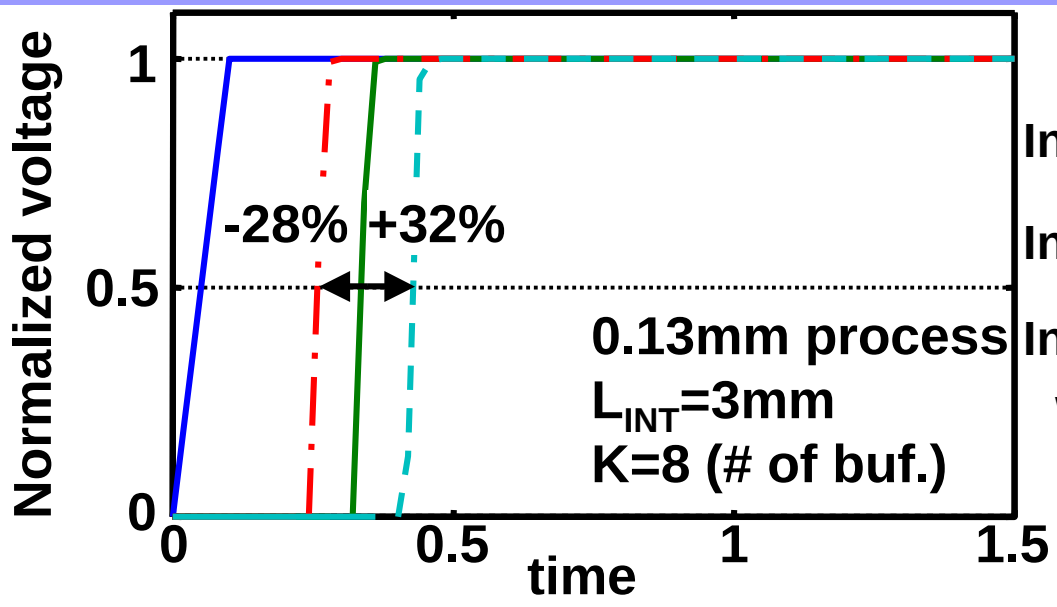


$$\frac{t_{pd}}{RC} = \frac{1}{2} + 2\eta - \frac{1}{\sqrt{6}} \log \frac{e}{2} \sqrt{1 + 8\eta + 6\eta^2}$$

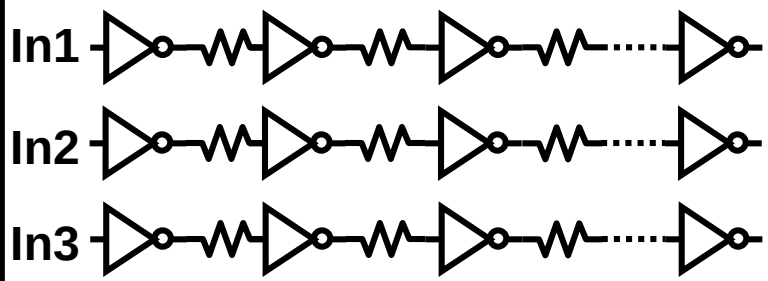
$$\approx 1.63\eta + 0.37 (\eta \le 2) \; (\eta = C_C / C)$$

H.Kawaguchi and T.Sakurai, "Delay and Noise Formulas for Capacitively Coupled Distributed RC Lines," 1998 ASPDAC, Digest of Tech. Papers, pp.35-43, Feb. 1998.

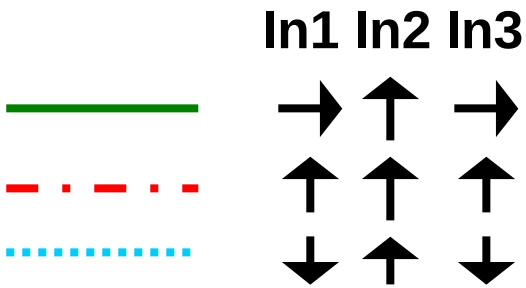
Coupling among Interconnections

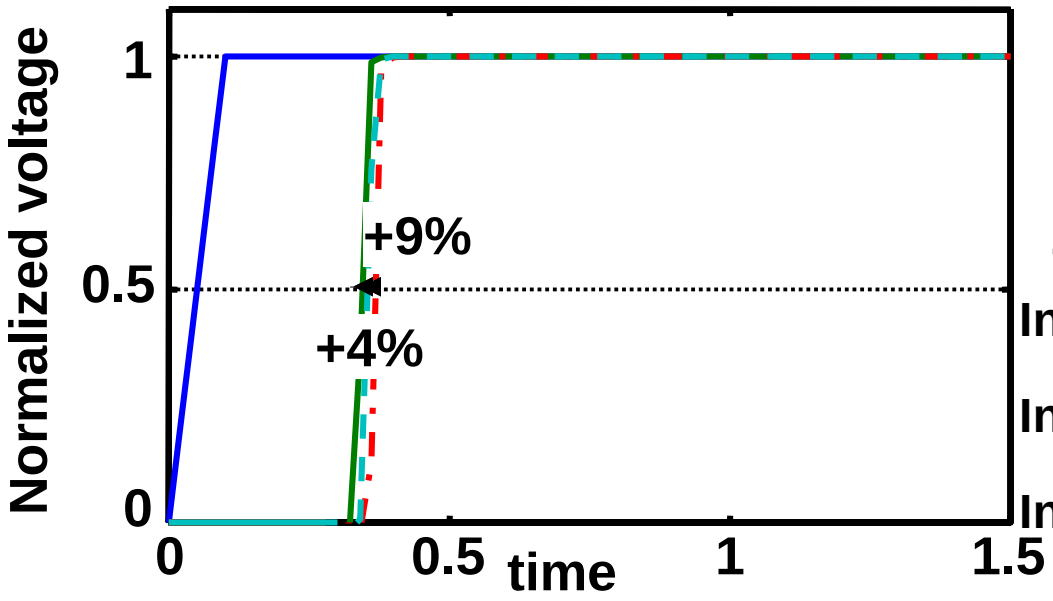


Normal buffer insertion

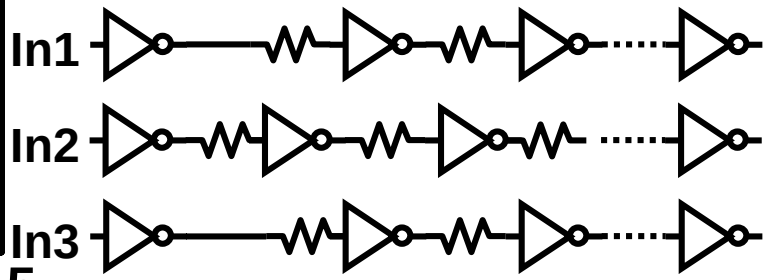


W_N/W_P=25μ/38μ

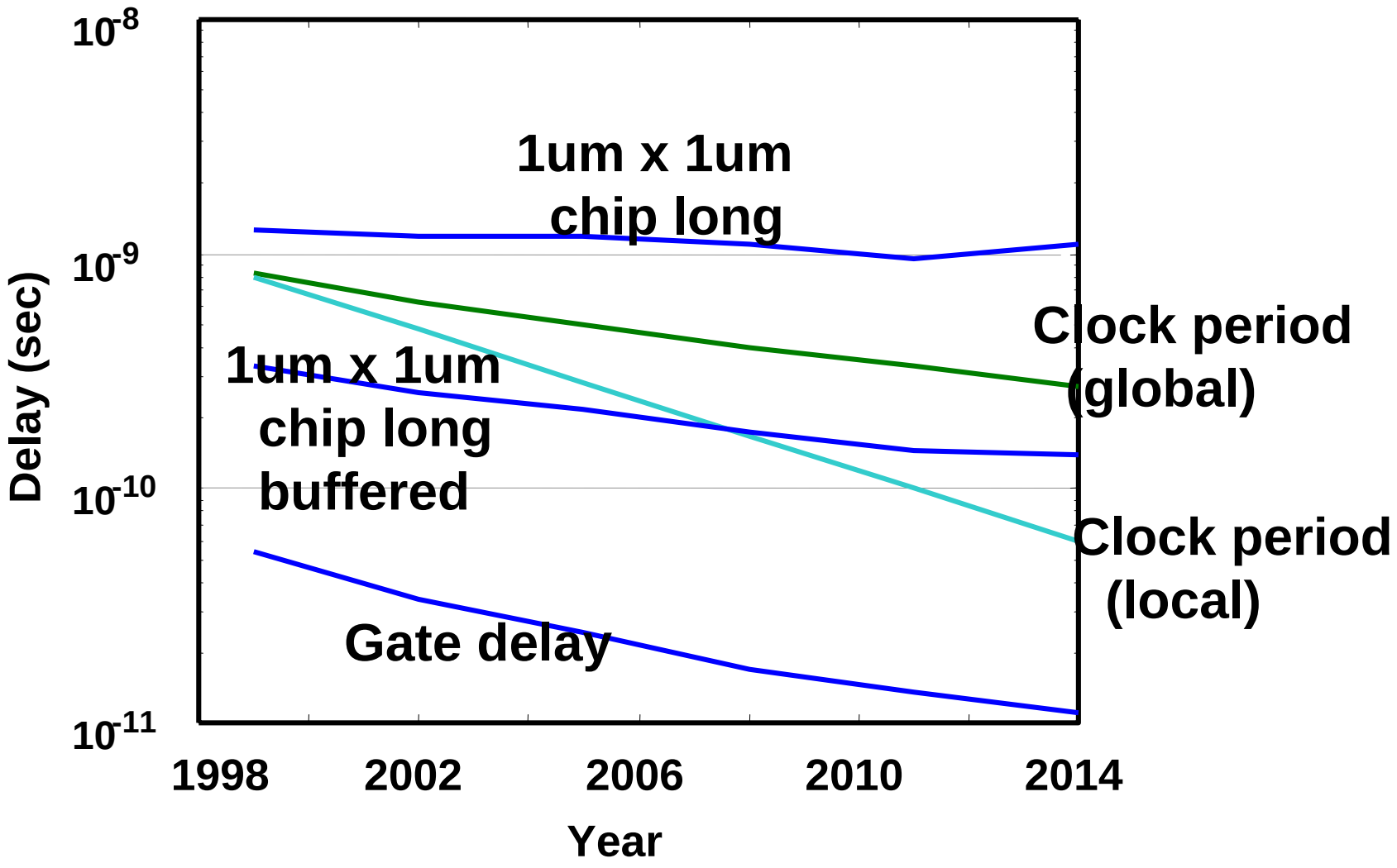




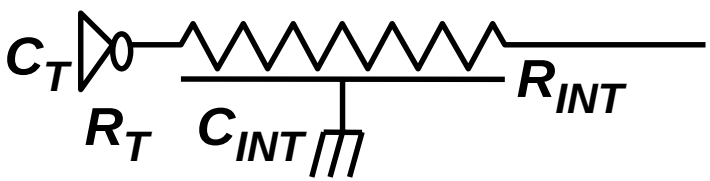
Zigzag buffer insertion



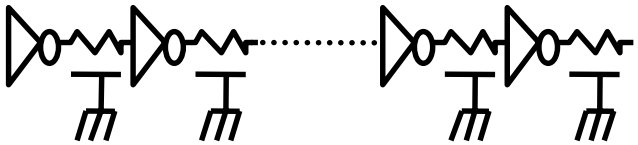
Buffered interconnect delay



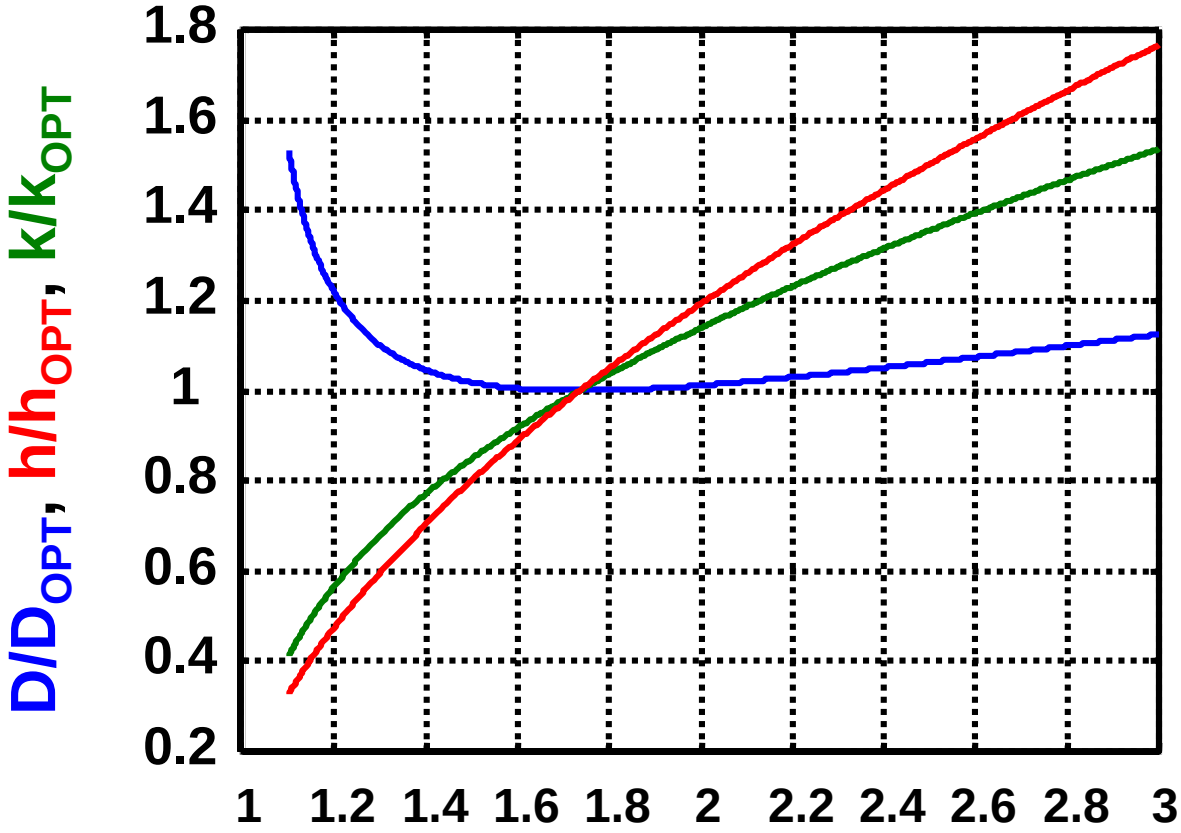
Power and delay optimization



a) Without repeaters



b) With repeaters



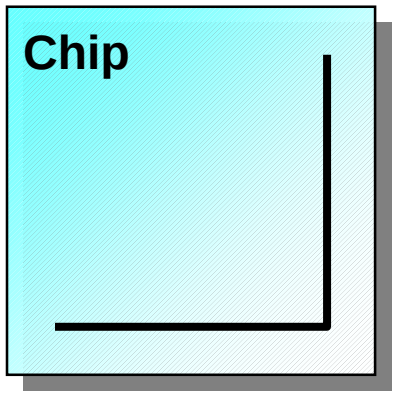
h: size of repeater

k: # of stages

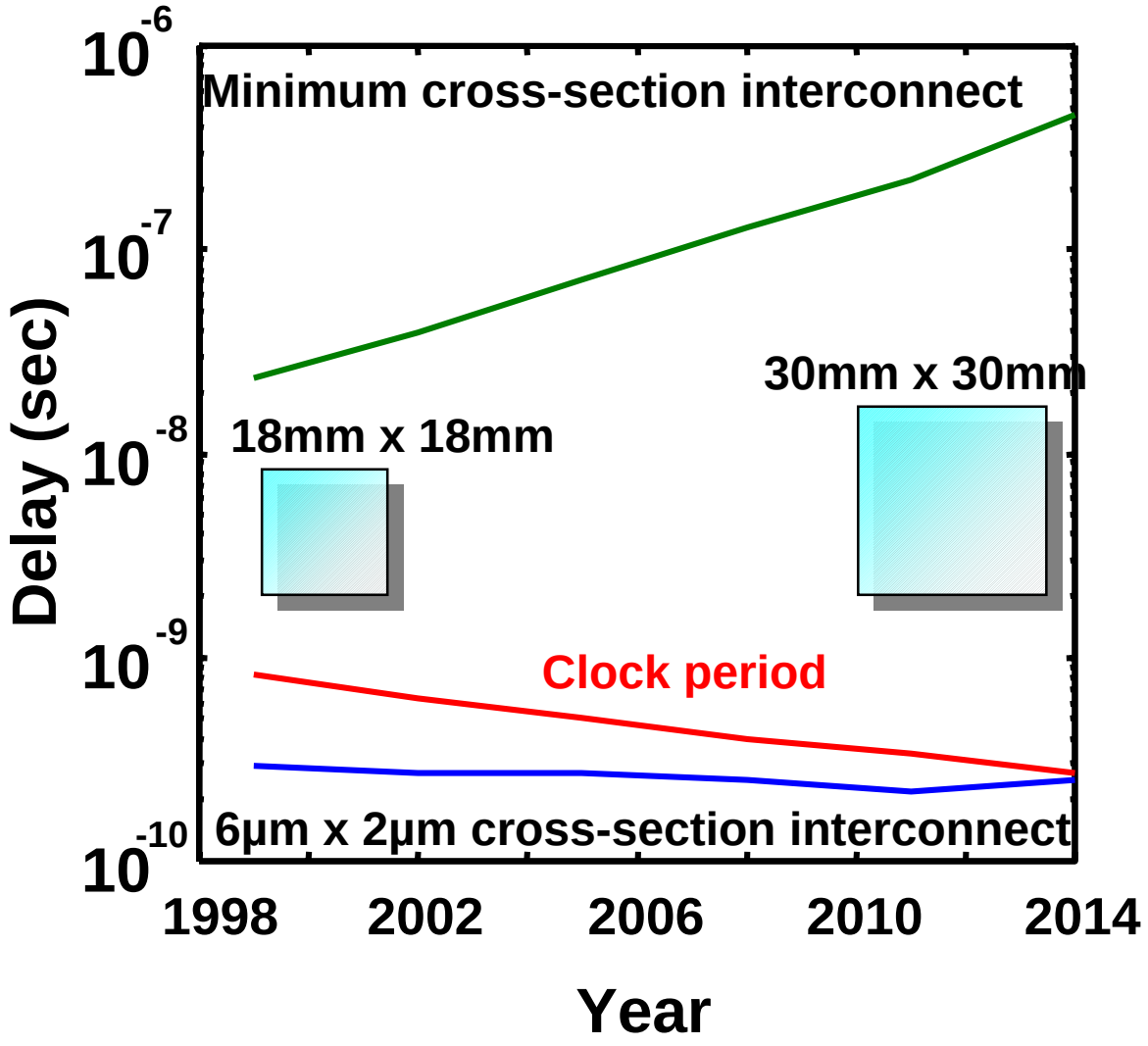
D: total delay

C_{INT}: interconnect capacitance

RC delay of global interconnections



Global interconnect



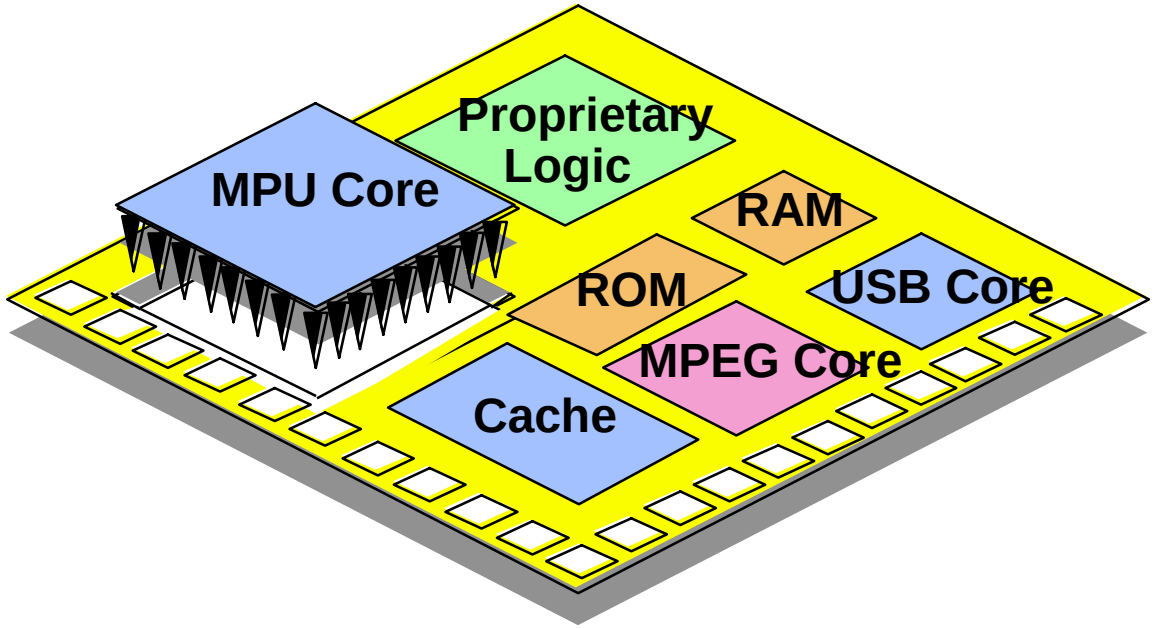
Three crises in VLSI designs

- Power crisis
- Interconnection crisis
- Complexity crisis

Overcome complexity crisis

- **Re-use and sharing of design**
- **Design in higher abstraction**

IP (A inc.)
IP (B univ.)
IP (C inst.)
IP (D semi.)

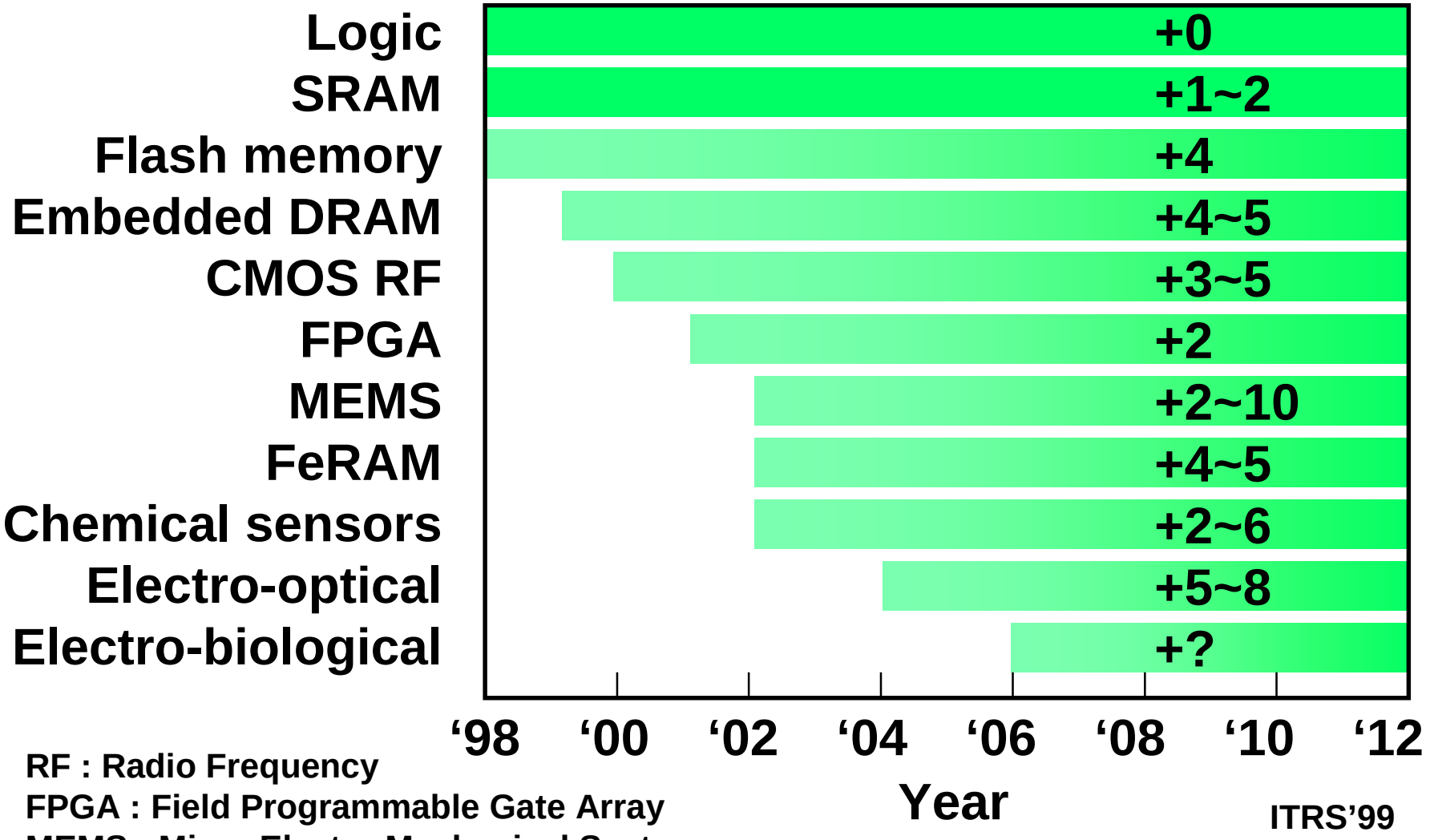


IP ; CPU, DSP, memories, analog, I/O, logic..
HW/FW/SW

Issues in System-on-Chip

- Un-distributed IP's (i.e. CPU, DSP of a certain company)
- Huge initial investment for masks & development
- IP testability, upfront IP test cost
- Process-dependent memory IP's
- Difficulty in high precision analog IP's due to noise
- Process incompatibility with non-Si materials and/or MEMS

Technologies integrated on a chip



ITRS’99

RF : Radio Frequency
FPGA : Field Programmable Gate Array
MEMS : Micro Electro Mechanical Systems
FeRAM : Ferroelectric RAM

System-in-Package

ELECTRONIC ENGINEERING
EE TIMES
eet.com The industry newspaper for engineers and technical management Monday, November 8, 1999

Expanding role of packaging seen relegating SoC to niche status

System-chip may topple . . .

By Robert Ristelhueber

INDIAN WELLS, CALIF. – The wheels might be coming off the system-on-chip (SoC) bandwagon, if the chatter at last week's Dataquest Semiconductor conference is any barometer of industry sentiment. Heavyweights including IBM and Lucent Technologies indicated that costs may relegate SoC to niche status, with new packaging techniques stepping into the breach.

"A couple of years ago we really thought that the embedded DRAM model would be the panacea for many applications," said John Kelly, general manager of IBM Microelectronics. "It's not always the right thing. In many applications it still remains much cheaper to do it with multichip modules. It gives you satisfactory performance and often for lower cost."

"We have systems-on-chip now that are really 'system on chips,' " said John Dickson, president of Lucent Technologies' Microelectronics Group. "We do it that way because it's most cost-effective, and the customer will prefer it that way because it offers more flexibility."

The subject was broached at the conference here by a Dataquest analyst who claimed that SoC designs will increasingly be supplanted in coming years by multichip packaging as higher mask costs squeeze SoC profitability.

Chip designers have often been willing to add mask steps

▶CONTINUED ON PAGE 6

IBM's Kelly: 'In many apps, cheaper to do it with multichip modules.'

. . . as industry grapples with impact of cores mode

By Peter Clarke and Brian Fuller

EDINBURGH, SCOTLAND – Intellectual property cores were a hot topic last week, both here at the IP99 Europe conference and at Dataquest Inc.'s annual semiconductor conference in Indian Wells, Calif. But as the industry struggles with new business models, new customer-supplier relationships and fast-moving technology, there was scant agreement on either side of the Atlantic on how the cores market will unfold.

IP-99

On one thing there was agreement: IP cores and design reus

In some apps, multichip modules do the job more cheaply, conference told

'System-in-package' could make SoC a niche

▶CONTINUED FROM PAGE 1

and complexity to their logic devices in order to place analog and memory functions onto chips. "But when we get below 0.2 micron we get a cost shock, and the [return on investment] will be diminished or even eliminated in many cases," said Clark Fuhs, vice president and director of Dataquest's Semiconductor Manufacturing Programs.

Mask costs will dramatically rise at deep submicron because of the use of phase-shift and optical proximity correction techniques as well as more expensive, 193-nm lithography equipment, putting low-volume SoC at a cost disadvantage, Fuhs said.

Militating against SoC designs for many applications is the wide disparity in revenue per square inch among the various blocks in the chip, Fuhs said. "The DSP or microprocessor block can be getting $150 or $200 per square inch, the FPGA about $120, the analog block about $35, the memory block about $50 to $60 . . . You're basically diluting your high-value logic pieces with all these other low-value pieces, yet you're adding cost because you're adding mask levels."

An alternative is to fabricate the different blocks as discrete chips, placed close together using chip-scale packaging, Fuhs said. "This enables you to build the pieces in fabs that are optimized for those pieces. You can build analog in a 0.7-micron fab, standard logic can be done in 0.35 or even 0.5 micron, and for the memory you can buy a wafer from somebody and break it up. The package is more expensive, but the overall system cost is going to be substantially less.

"The concept here is to take some level of interconnect . . . and simply move [it] from the chip into the package."

Fuhs noted that Intel's Pentium III is actually an 11-level-metal device—six levels of aluminum inside the chip and five levels of copper outside. And he showed a photograph of a Sony digital Handycam, which he said contains 20 chip-scale devices, "so this technology is here, it's real."

National's Halla touts 'integrated disintegration.'

In the not-too-distant future, he said, wafer foundries will give customers a choice of implementing a design either as a system-on-chip or as several discrete devices using chip-scale packaging.

To survive, the SoC must evolve to fit a more standard-product model that would allow it to increase volume and become more cost-efficient, Fuhs said. He predicted that within five years, multichip packaging will be growing faster than SoC designs.

That view has its detractors. "Mask sets cost in excess of a couple hundred thousand dollars, whether you do small chips or large chips," said National Semiconductor Corp. chief executive officer Brian Halla, who has championed the notion of an information appliance-on-a-chip. "I can get tremendously more performance out of the same square inches of silicon by having it all together instead of having it two inches apart on a board.

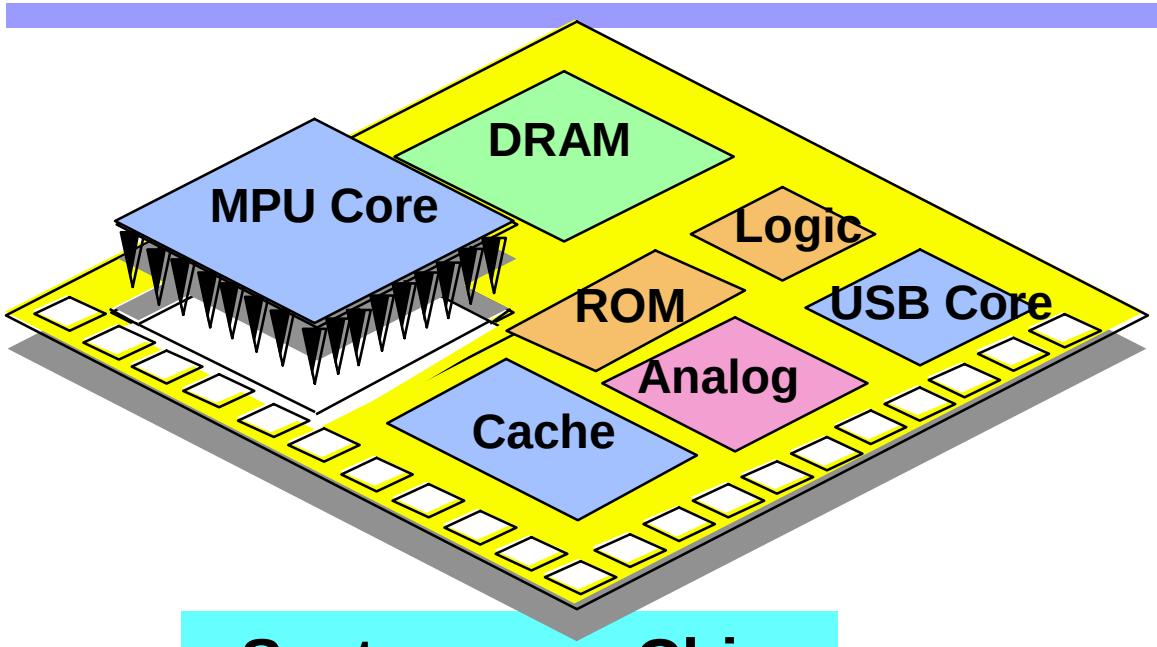
"SoC isn't a marketing crusade anymore; it's something you can do because the technology allows it," Halla added. "A very small die can contain an awful lot of functionality."

Halla noted that Intel used to say graphics shouldn't be combined with the microprocessor, because the pace of innovation differs between those parts; but Intel's upcoming Timna processor, he said, combines both functions.

"Having said all that, there are cases where we agree [about putting a system on a package]," he said. "There is a substrategy of ours called integrated disintegration, which means there are analog functions you can pull off the chip because they are such a tiny portion of the overall chip, and yet they are the most difficult thing to port to the next-generation [process] technology."

IBM's Kelly said that "SoC integration has to be done se-

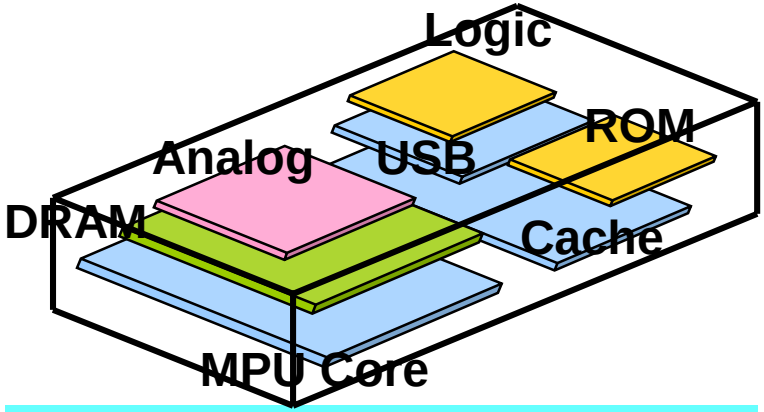
SoC vs. SiP



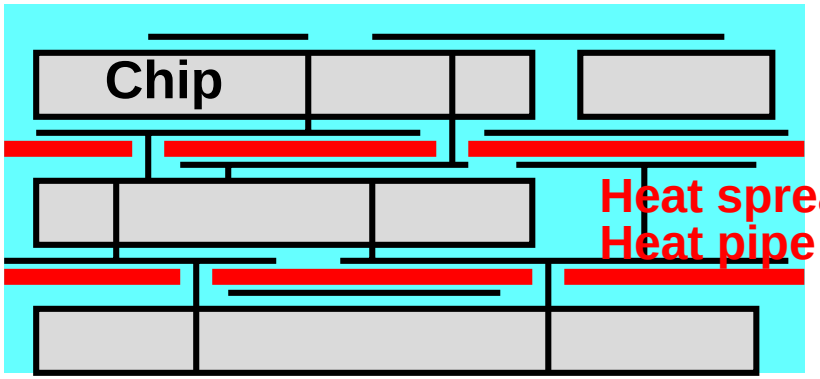
System on a Chip

- Smaller area
- Shorter interconnect
- Optimized process for each die (Analog, DRAM, MEMS…)
- Good electrical isolation
- Through-chip via

- Heat dissipation is an issue



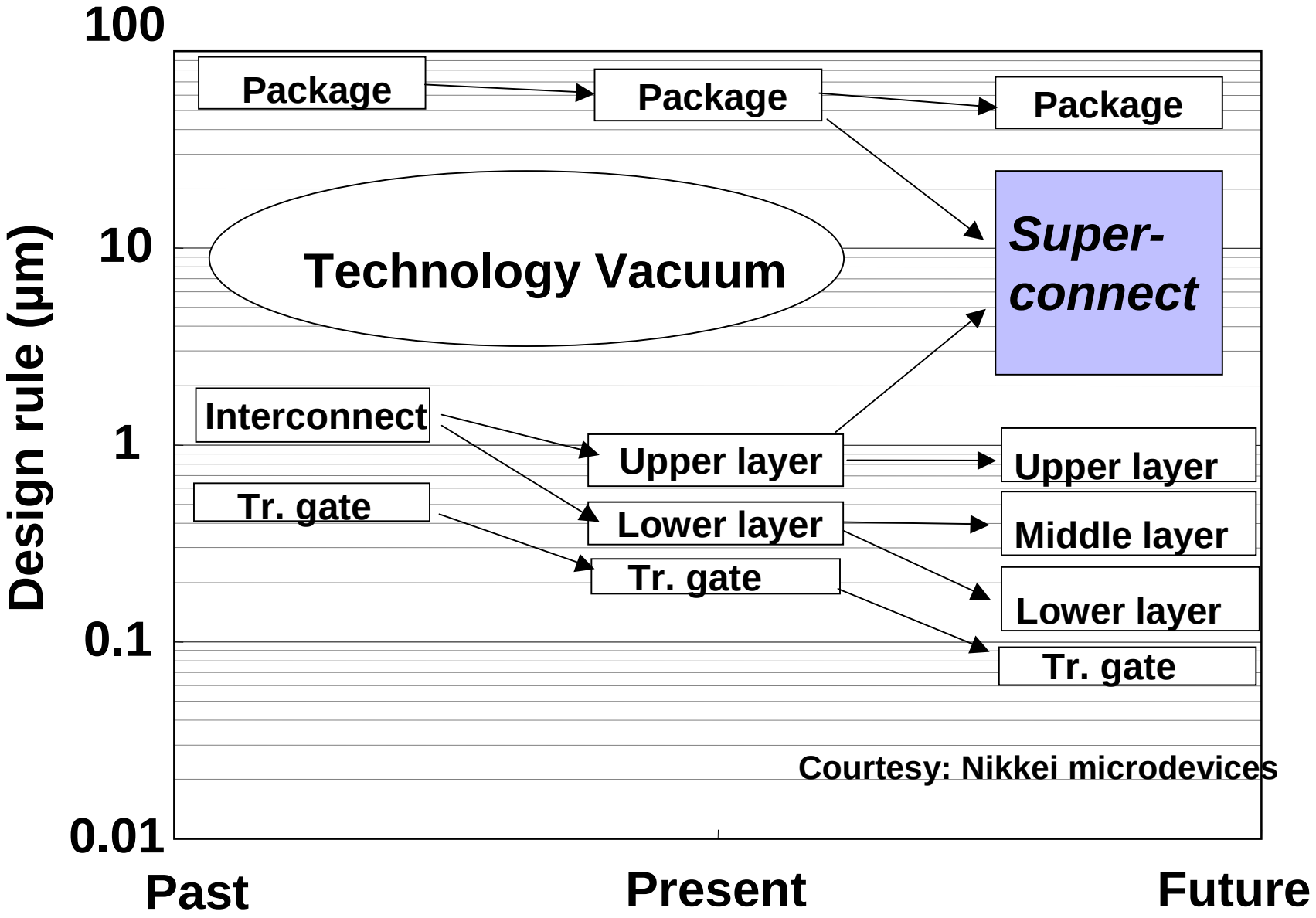
System in a Package



Issues in System-in-Package

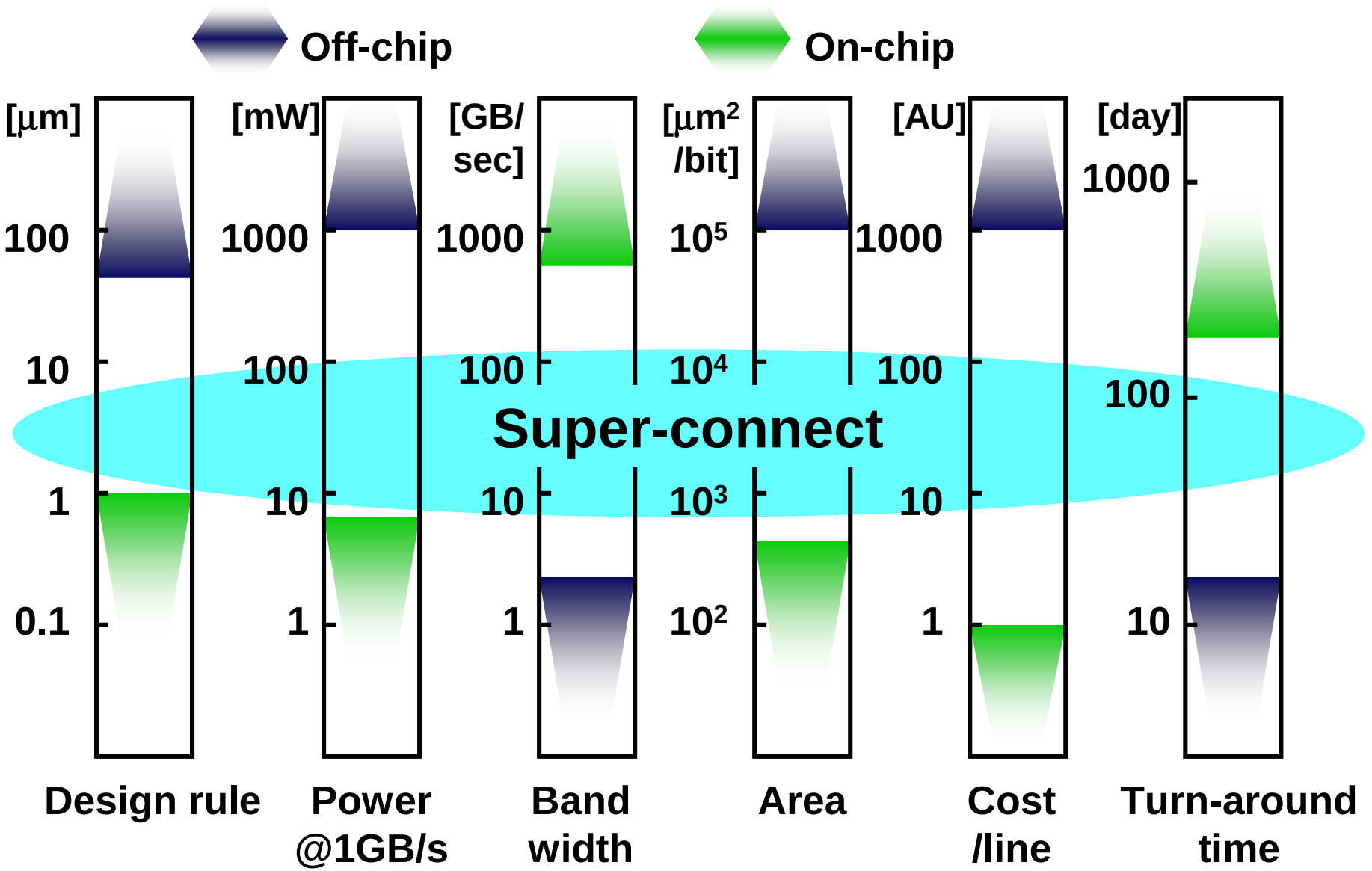
- **Special design tools for placement & route for co-design of LSI's and assembly**
- **High-density reliable substrate and metallization technology**
- **low-cost, available known good die (reworkablility and module testing)**

Super-connect technology



T.Sakurai

Super-connect



T.Sakurai

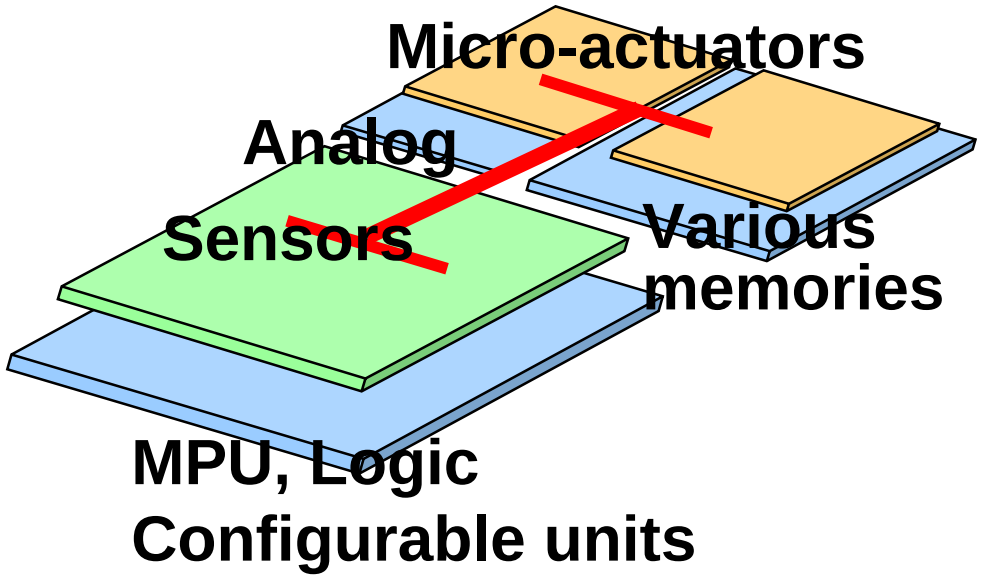
LSI in 2014

Year	Unit	1999	2014	Factor
Design rule	μm	0.18	0.035	0.2
Tr. Density	/cm2	6.2M	390M	30
Chip size	mm2	340	900	2.6
Tr. Count per chip (μP)		21M	3.6G	170
DRAM capacity		1G	1T	1000
Local clock on a chip	Hz	1.2G	17G	14
Global clock on a chip	Hz	1.2G	3.7G	3.1
Power	W	90	183	2.0
Supply voltage	V	1.5	0.37	0.2
Current	A	60	494.6	8
Interconnection levels		6	10	1.7
Mask count		22	28	1.3
Cost / tr. (packaged)	μcents	1735	22	0.01
Chip to board clock	Hz	500M	1.5G	3.0
# of package pins		810	2700	3.3
Package cost	cents/pin	1.61	0.75	0.5

International Technology Roadmap for Semiconductors 1998 update sponsored by the Semiconductor Industry Association in cooperation with European Electronic Component Association (EECA) , Electronic Industries Association of Japan (EIAJ), Korea Semiconductor Industry Association (KSIA), and Taiwan Semiconductor Industry Association (TSIA) , International Technology Roadmap for Semiconductors: 1999 edition. Austin, TX:International SEMATECH, 1999.

T.Sakurai

Possible electronic system in 2014



- Sensors/actutors
- 0.035μm 3.6G Si FET's with VTH & VDD control
- Locally synchronous 17GHz clock, globally asynchronous
- Chip / Package / Board system co-design for power lines, clocks, and long wires (super-connect)

Summary – Interconnect & SiP

- New possibilities with buffered interconnects may open up new tools opportunity.
- Silicon-in-Package needs new tools that support co-design of VLSI's, package and assembly.

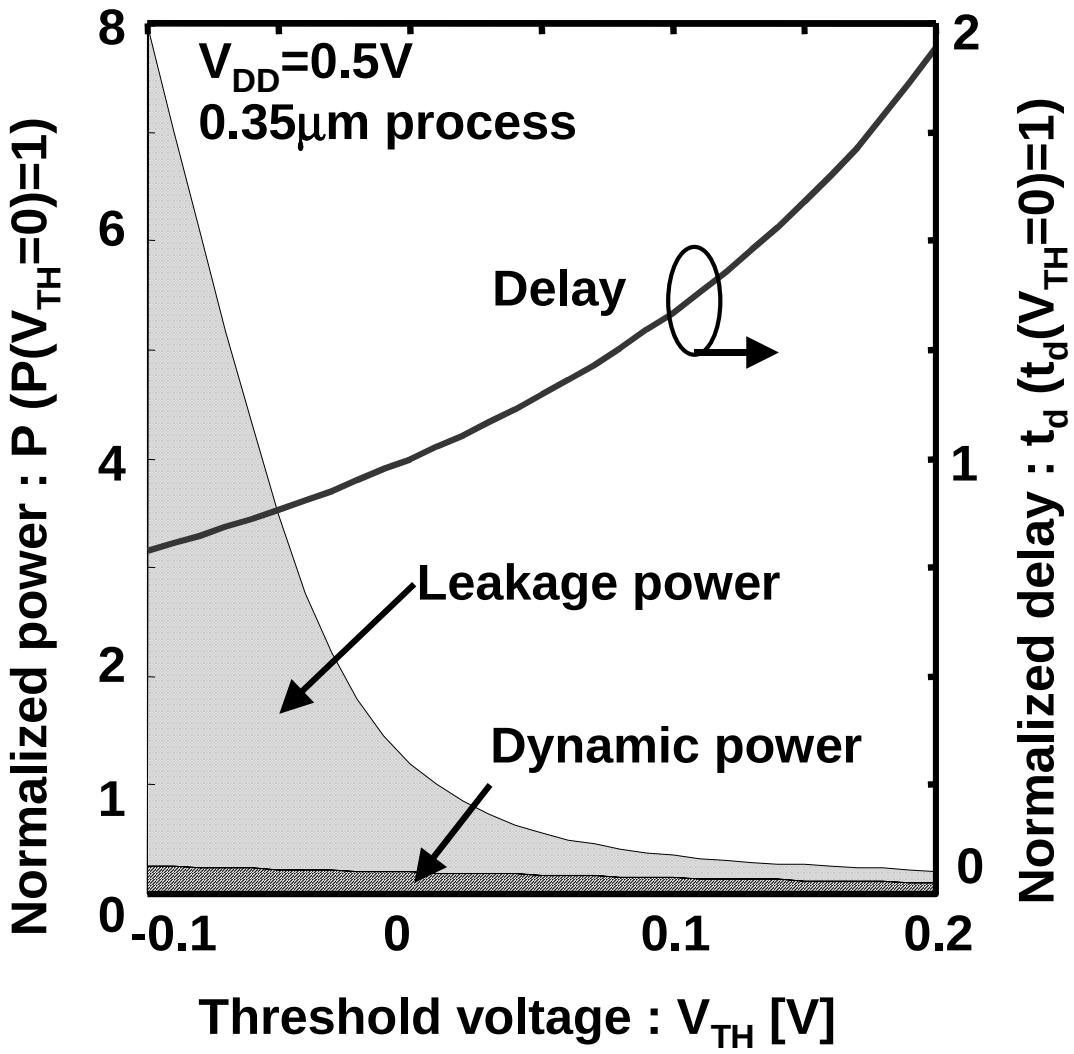
  New design closure issues

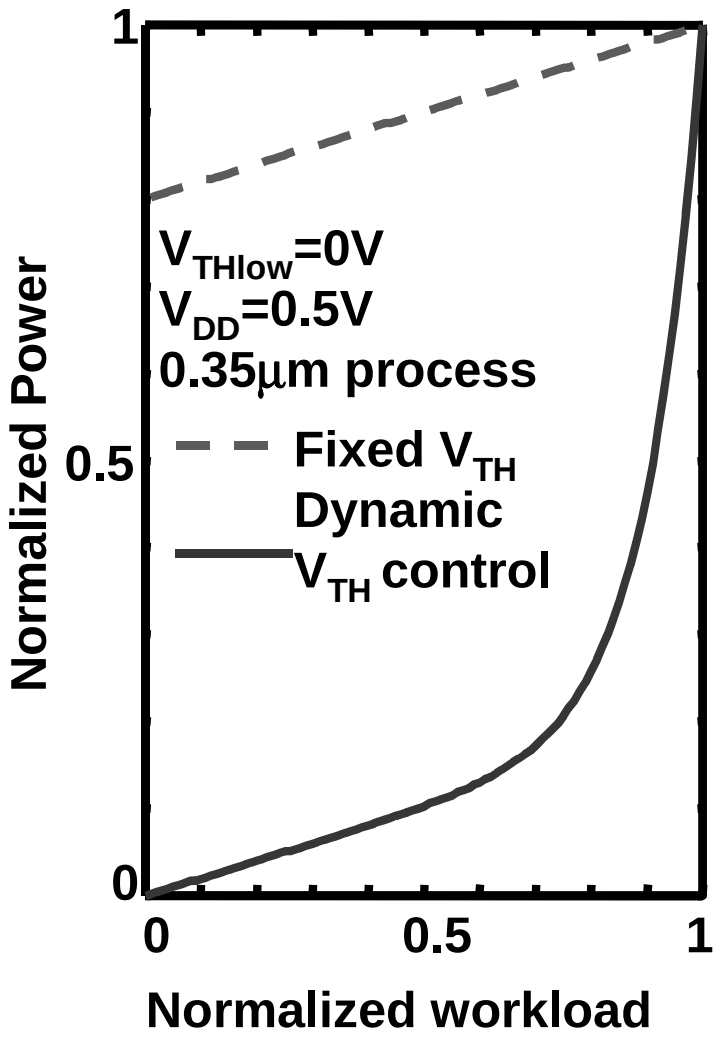
Controlling V_{DD} and V_{TH} for low power

	Active	Stand-by
Multiple V_{TH}	Dual-V_{TH}	MTCMOS
Variable V_{TH}	V_{TH} hopping	VTCMOS
Multiple V_{DD}	Dual-V_{DD}	Boosted gate MOS
Variable V_{DD}	V_{DD} hopping	

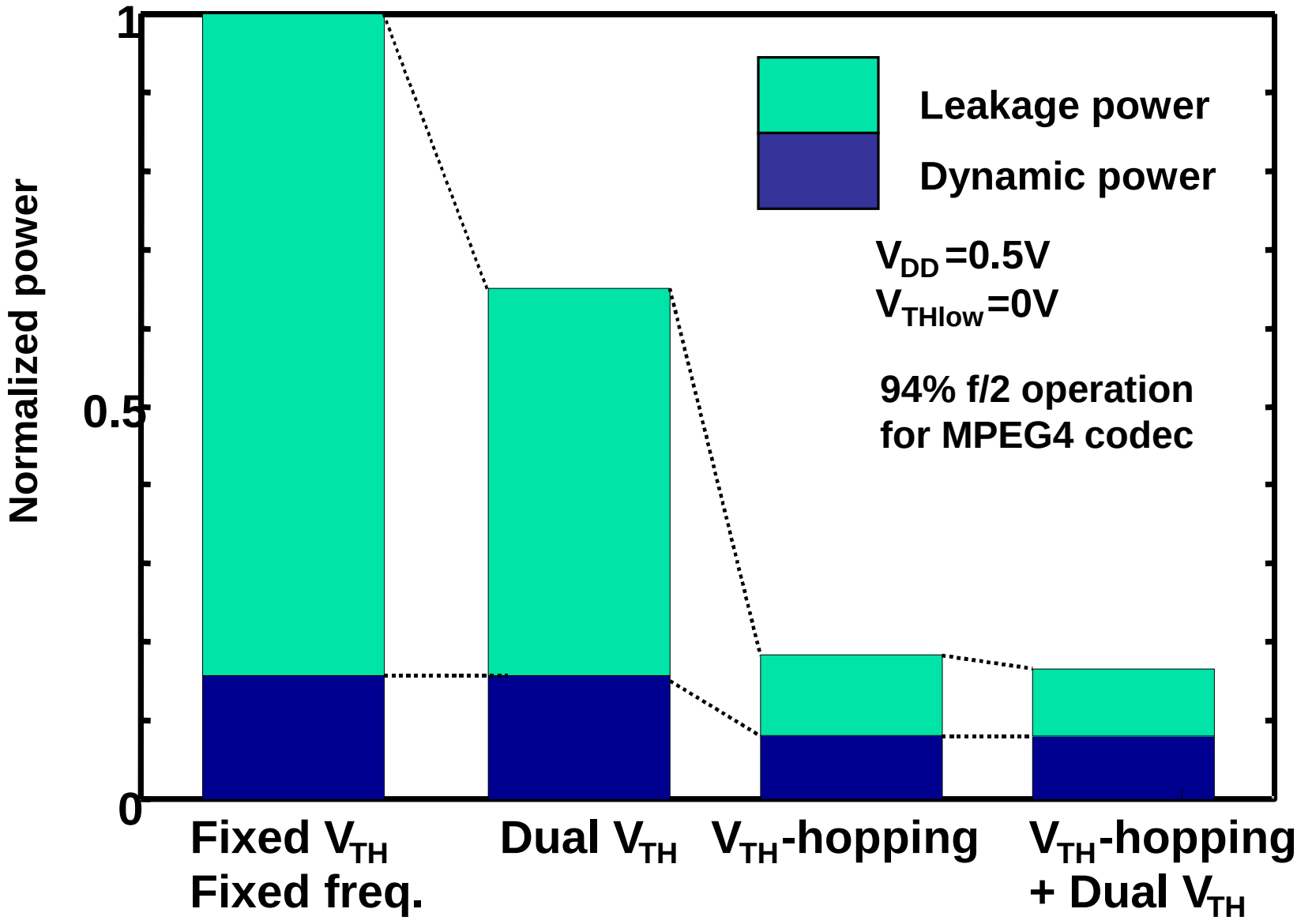
K. Nose, M.Hirabayashi, H.Kawaguchi, S.Lee and T.Sakurai, "VTH-hopping Scheme for 82% Power Saving in Low-voltage Processors," to be published, CICC 2001.

V_{TH} hopping


V_{DD}=0.5V
0.35μm process
Delay
Leakage power
Dynamic power
Normalized power : P (P(V_{TH}=0)=1)
Normalized delay : t_d (t_d(V_{TH}=0)=1)
Threshold voltage : V_{TH} [V]
8
6
4
2
1
0
2
1
0
-0.1
0
0.1
0.2


V_{THlow}=0V
V_{DD}=0.5V
0.35μm process
Fixed V_{TH}
Dynamic
V_{TH} control
Normalized Power
Normalized workload
1
0.5
0
0
0.5
1

V_{TH} hopping


Normalized power
1
0.5
0
Leakage power
Dynamic power
V_{DD}=0.5V
V_{THlow}=0V
94% f/2 operation
for MPEG4 codec
Fixed V_{TH}
Fixed freq.
Dual V_{TH}
V_{TH}-hopping
V_{TH}-hopping
+ Dual V_{TH}

Other important technologies for low-power

$$P = \alpha f C Vs V_{DD} + \text{leak power}$$

Low-C

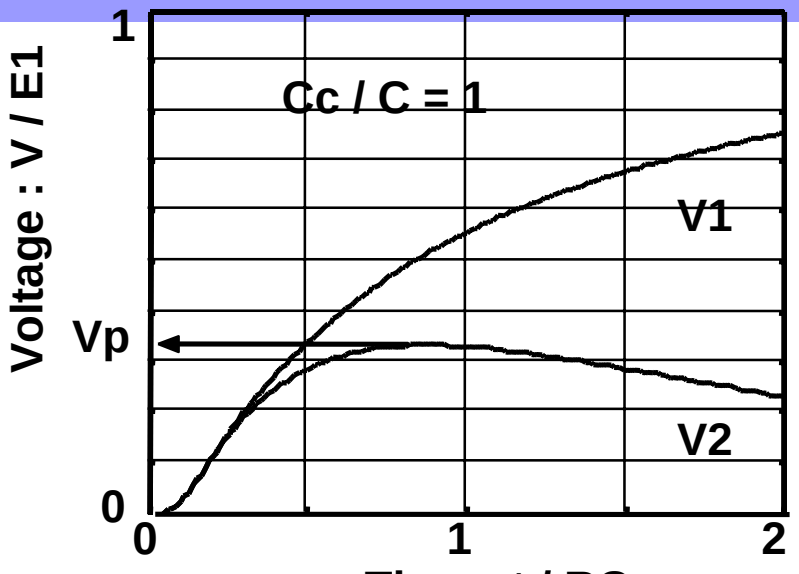
- Gate-sizing, low-power cell library
- Memory embedding
- Mixed digital-analog

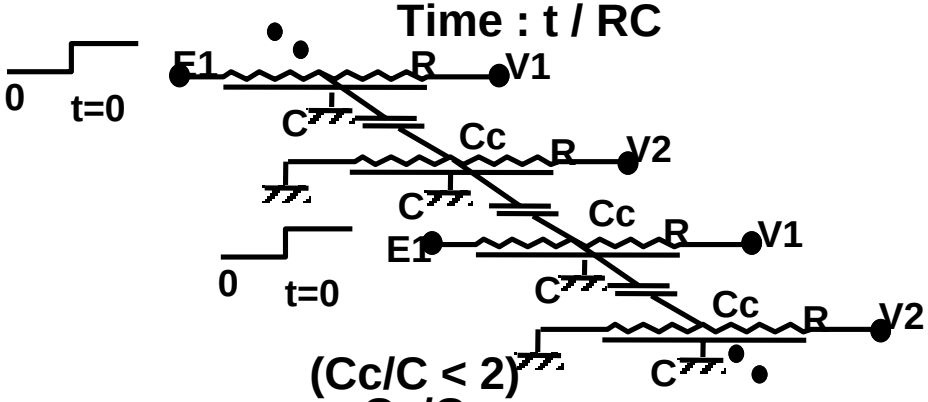
Low- αf

- Multi-f
- gated clock

System-level power prediction/optimizatiion

Coupling noise in RC bus

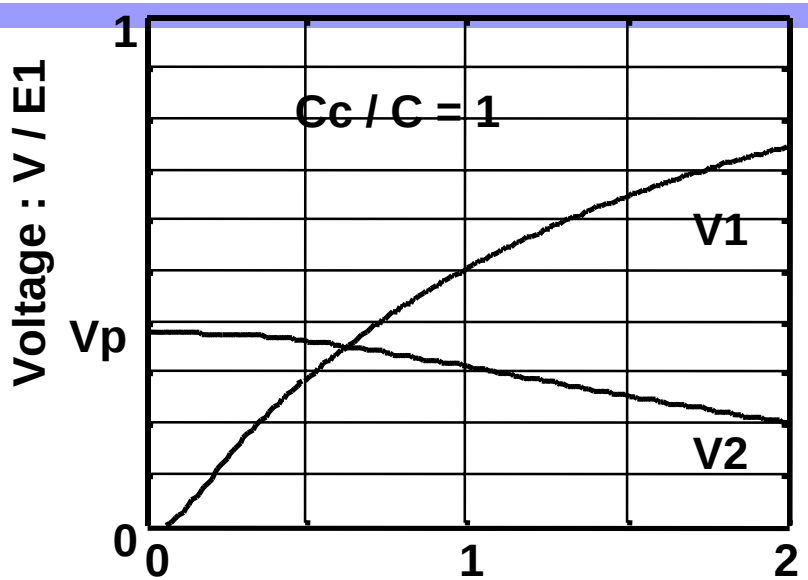


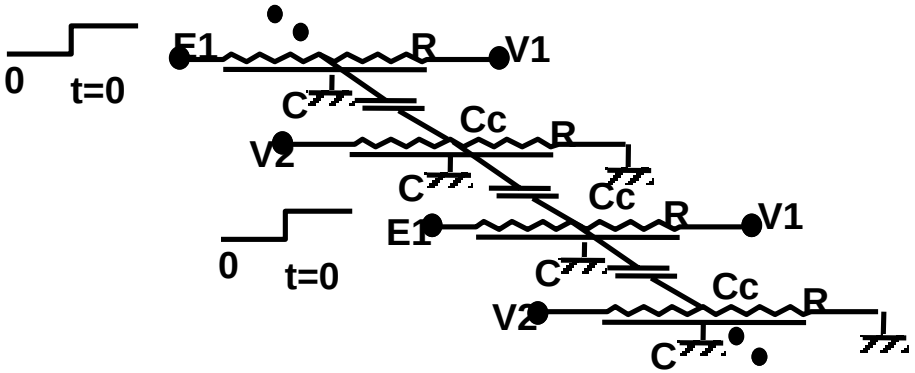


(Cc/C < 2)

$$Vp \approx \frac{Cc/C}{1+2Cc/C} \quad \text{(Bus)}$$

$$Vp \approx \frac{2Cc/C}{2+3Cc/C} \quad \text{(Three lines)}$$

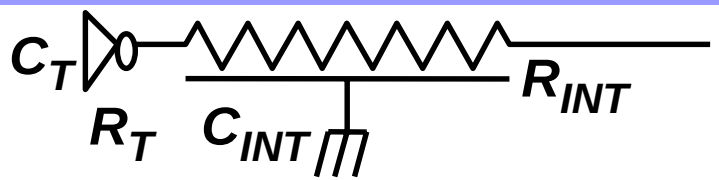




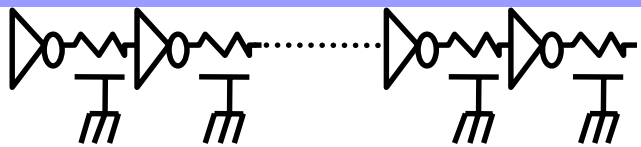
$$Vp = \frac{\sqrt{1+4Cc/C}-1}{\sqrt{1+4Cc/C}+1}$$

H.Kawaguchi and T.Sakurai, "Delay and Noise Formulas for Capacitively Coupled Distributed RC Lines," ASPDAC, Digest of Tech. Papers, pp.35-43, Feb. 1998.

Repeaters



a) Without repeaters



b) With repeaters

$$t_{05} \approx 0.377 R_{INT} C_{INT} + 0.693 (R_T C_T + R_T C_{INT} + R_{INT} C_T)$$

C_0 : Gate capacitance of minimum MOSFET

R_0 : Gate effective resistance of minimum MOSFET

$$Delay \approx k \left[p_1 \frac{R_{INT}}{k} \frac{C_{INT}}{k} + p_2 \left(\frac{R_0}{h} h C_0 + \frac{R_0}{h} \frac{C_{INT}}{k} + \frac{R_{INT}}{k} h C_0 \right) \right] : \text{Buffered}$$

$$\frac{\partial Delay}{\partial h} = 0 \rightarrow h_{OPT} = \sqrt{\frac{C_{INT} R_0}{R_{INT} C_0}} : \text{Optimized size of buffer inverter}$$

$$\frac{\partial Delay}{\partial k} = 0 \rightarrow k_{OPT} = \sqrt{\frac{p_1}{p_2}} \sqrt{\frac{R_{INT} C_{INT}}{R_0 C_0}} : \text{Optimized number of stages}$$

$$Delay_{OPT} = 2 \left(\sqrt{p_1 p_2} + p_2 \right) \sqrt{R_{INT} C_{INT} R_0 C_0} \approx 2.4 \sqrt{\tau_{INT} \tau_{MOS}}$$

$$Cap.\ of\ gates = k_{OPT} h_{OPT} C_0 = \sqrt{p_1 / p_2} C_{INT} = 0.73 C_{INT}$$

Necessity for Low-Power Design

Power range	Concerns	Typical applications (All need high-perf.)
< 0.1W	• Battery life	Portable • PDA • Communications
~ 1W	• Inexpensive package limit • System heat (10W / box)	Consumer • Set-Top-Box • Audio-Visual
> 10W	• Ceramic package limit • IR drop of power lines	Processor • High-end MPU's • Multimedia DSP's

VLSI Design in 2010

Designing a map of 10m wide roads for a world atlas

T.Sakurai