セミコン関西2000 ULSI技術セミナー　’00/6/2

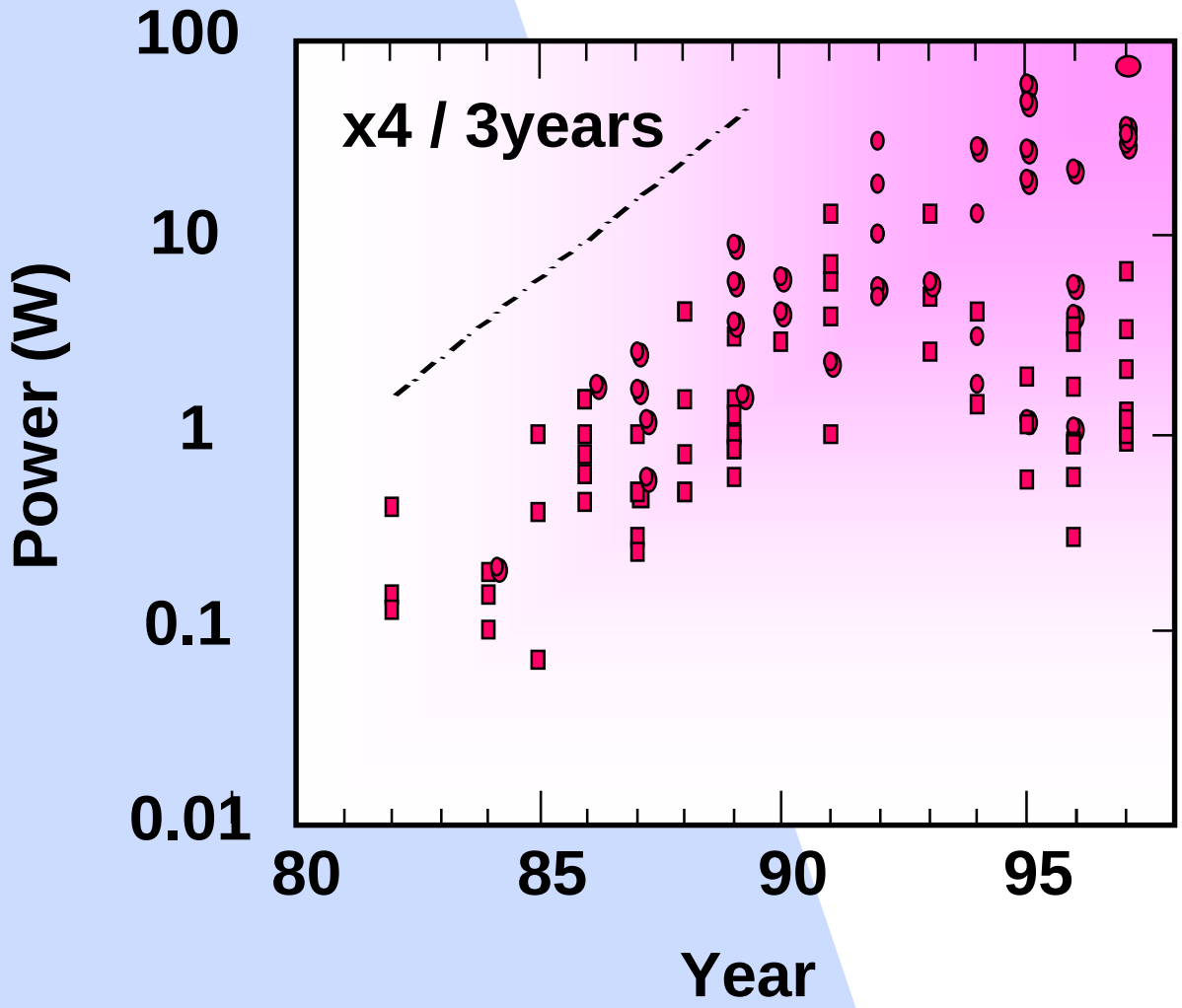
設計からみた低消費電力・高速化技術

東京大学、国際・産学共同研究センター、生産技術研究所
桜井貴康
E-mail:tsakurai@iis.u-tokyo.ac.jp

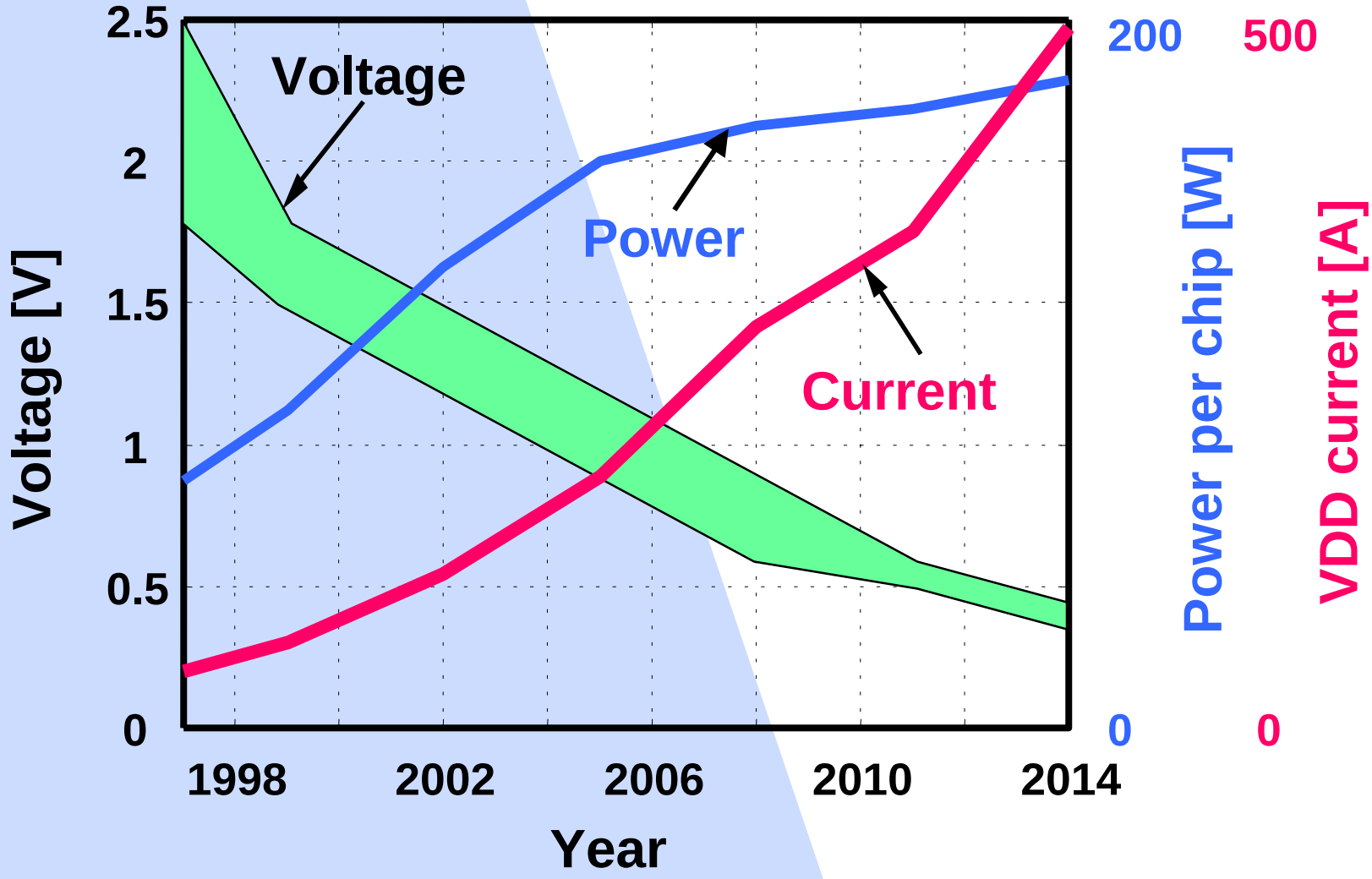
- Device / circuit collaboration: dual-oxide BGMOS
- Circuit / software collaboration: voltage hopping
- Architectural approaches

Ever increasing VLSI power

(Power consumption of processors published in ISSCC)



V_{DD}, power and current trend



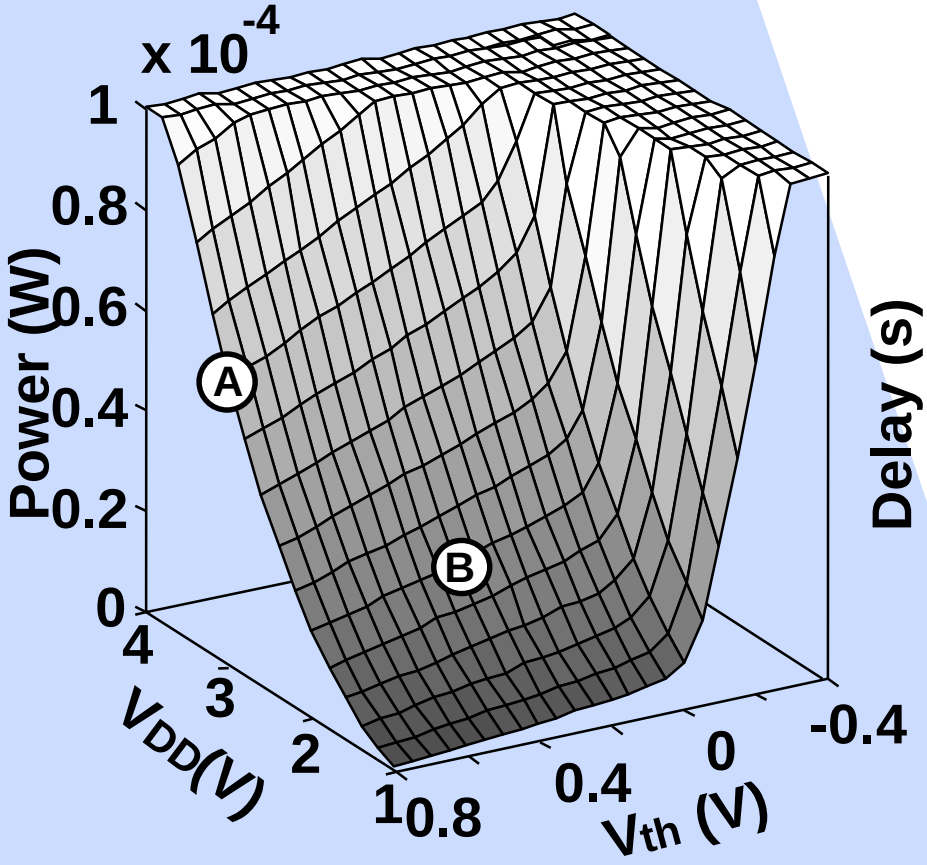
International Technology Roadmap for Semiconductors 1998 update sponsored by the Semiconductor Industry Association in cooperation with European Electronic Component Association (EECA) , Electronic Industries Association of Japan (EIAJ), Korea Semiconductor Industry Association (KSIA), and Taiwan Semiconductor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

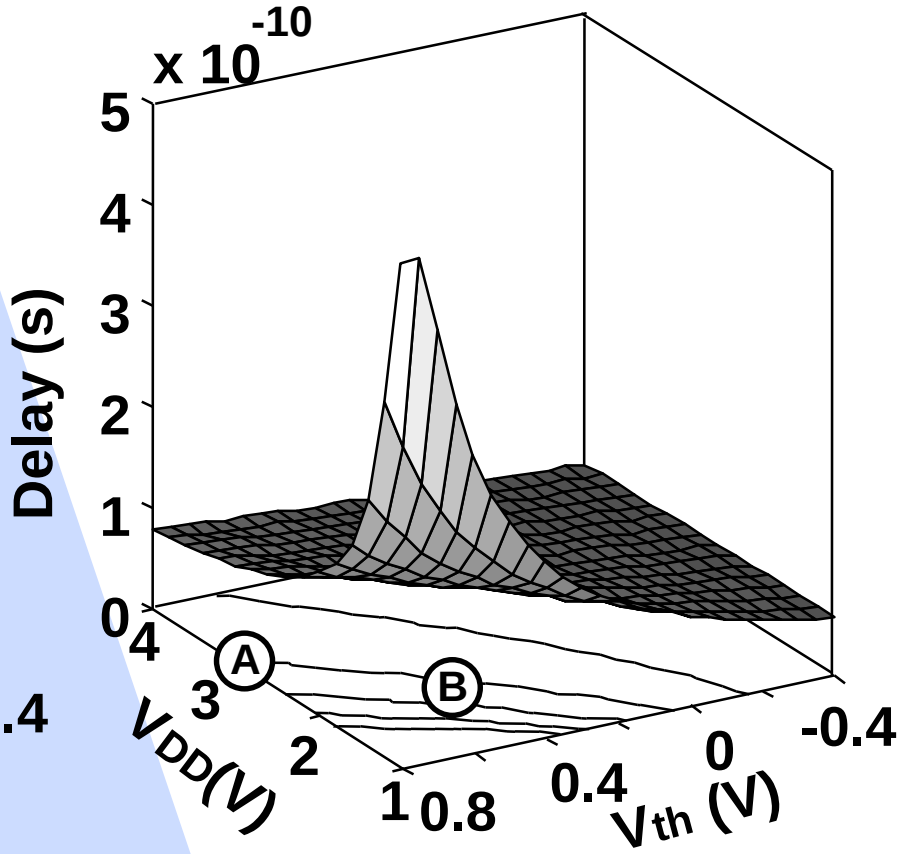
T.Sakurai

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



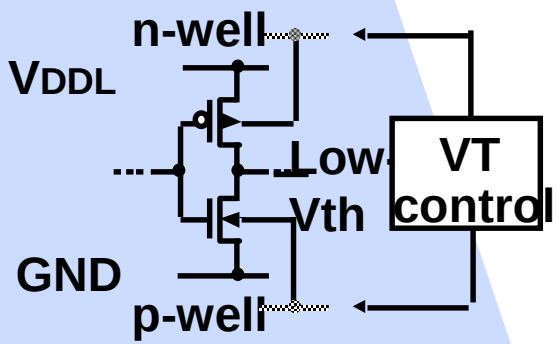
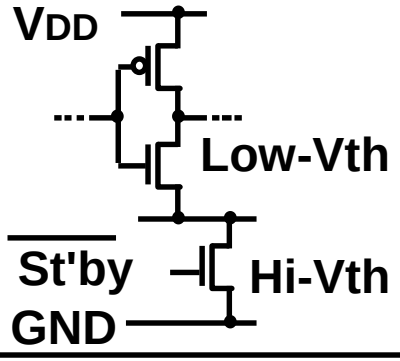


Controlling V_{DD} and V_{TH} for low power

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

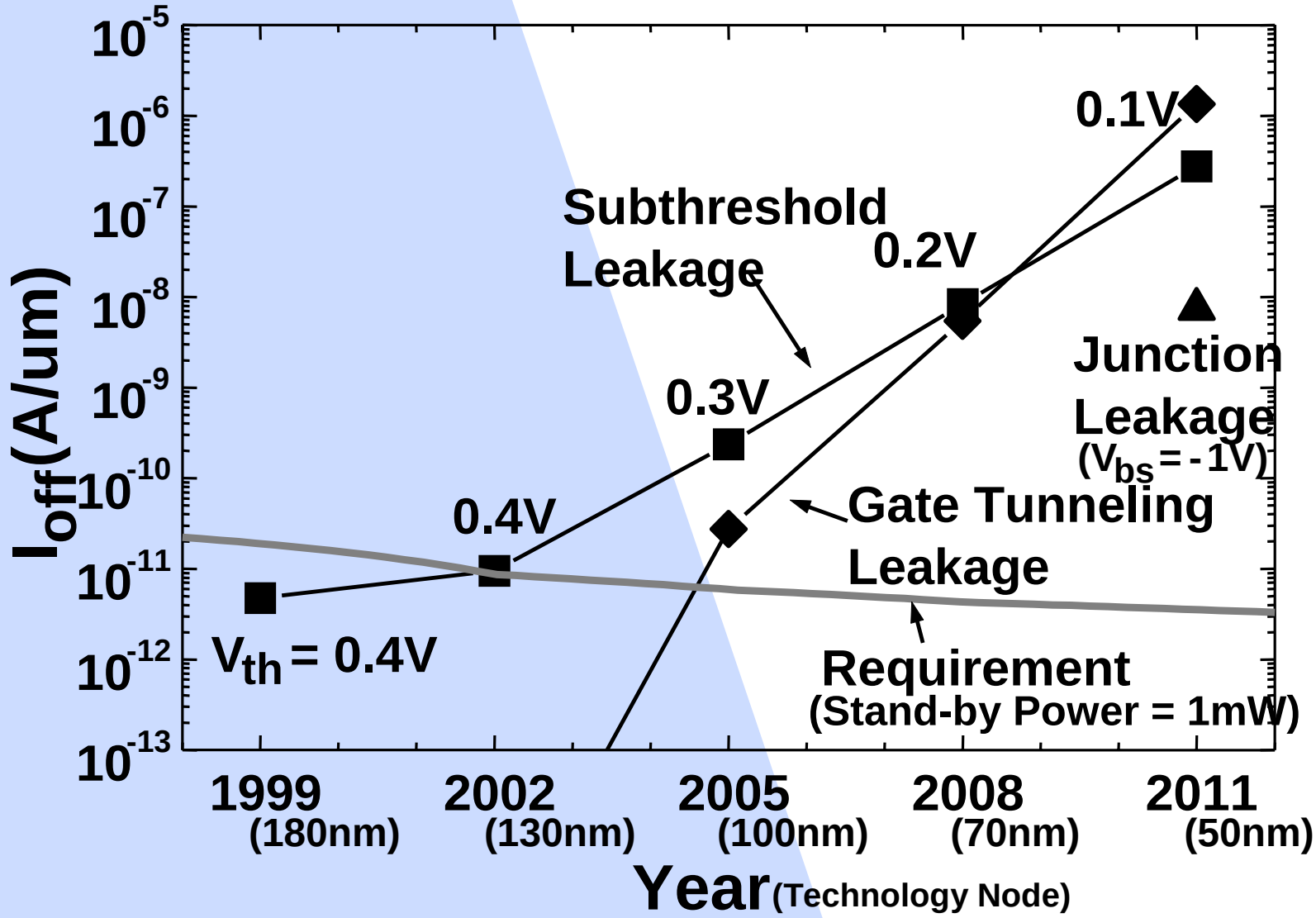
- **Multiple V_{TH}**
  - **Dual-V_{TH}, Multi-Threshold CMOS**
- **Variable V_{TH}**
  - **VTCMOS, Substrate bias control**
- **Multiple V_{DD}**
  - **Boosted gate MOS, Dual oxide/dual V_{DD}**
- **Variable V_{DD}**
  - **Voltage hopping, Software control**

Variable V_{TH} and Multi-V_{TH}

	Variable VTH	Multi-VTH
Principle		
	Threshold control with sub-bias	On-off control of internal VDD/VSS
Merit/ Demerit	+ Low leakage in standby + Already productized + Compensate Vth fluctuation + I_{DDQ} test + No serial MOSFET + Conventional design env. + Conventional F/F's - Triple well is desirable - Scalability? (junction leakage)	+ Low leakage in standby + Already productized - Compensate Vth fluctuation - I_{DDQ} test - Large serial MOSFET slower, larger, lower yield... + Conventional design env. - Special F/F's, Two V_{TH}'s - Ultra-low voltage region? - Delay fluctuates on activity

T.Sakurai

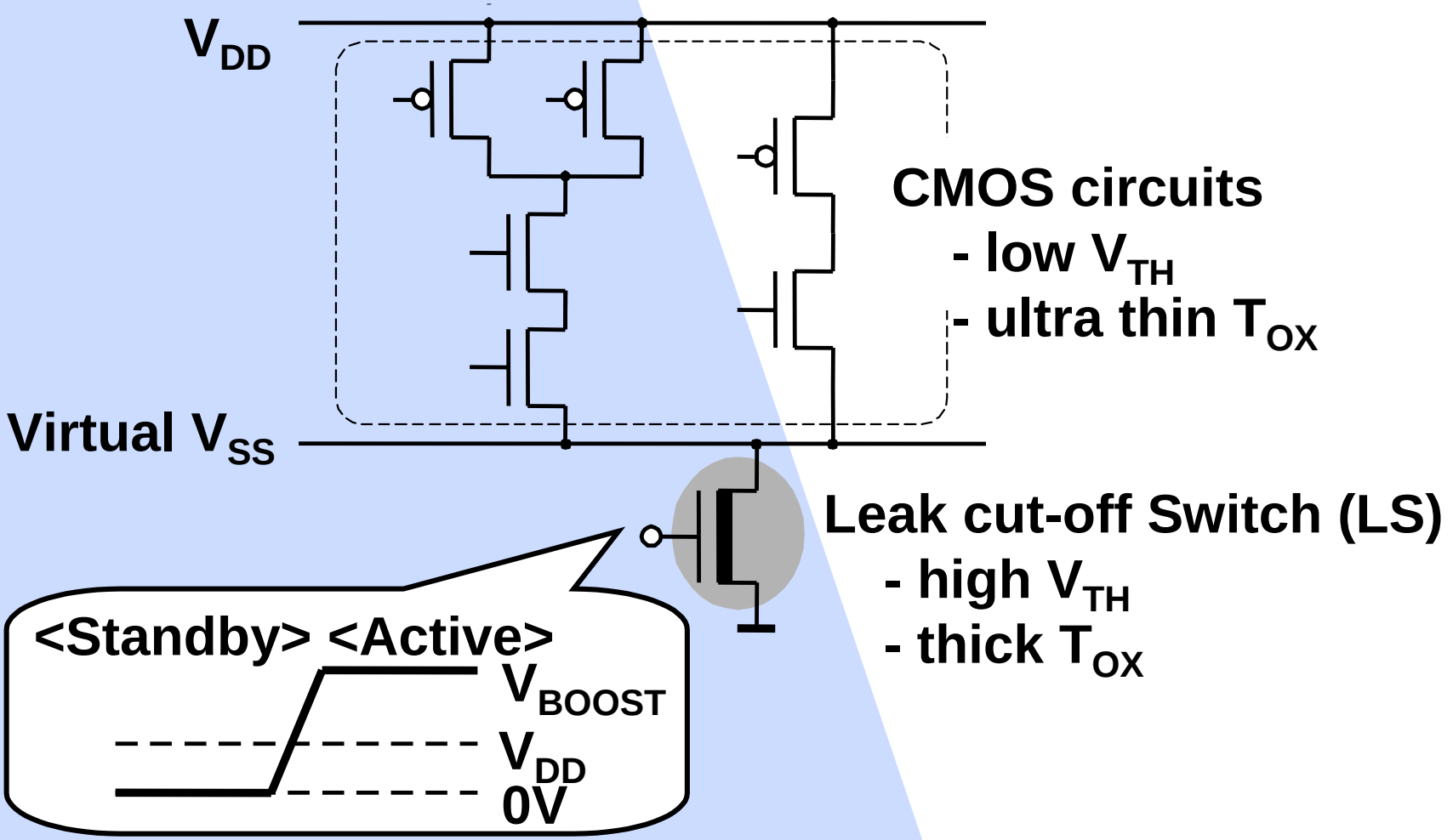
Transistors go leaky



T.Sakurai

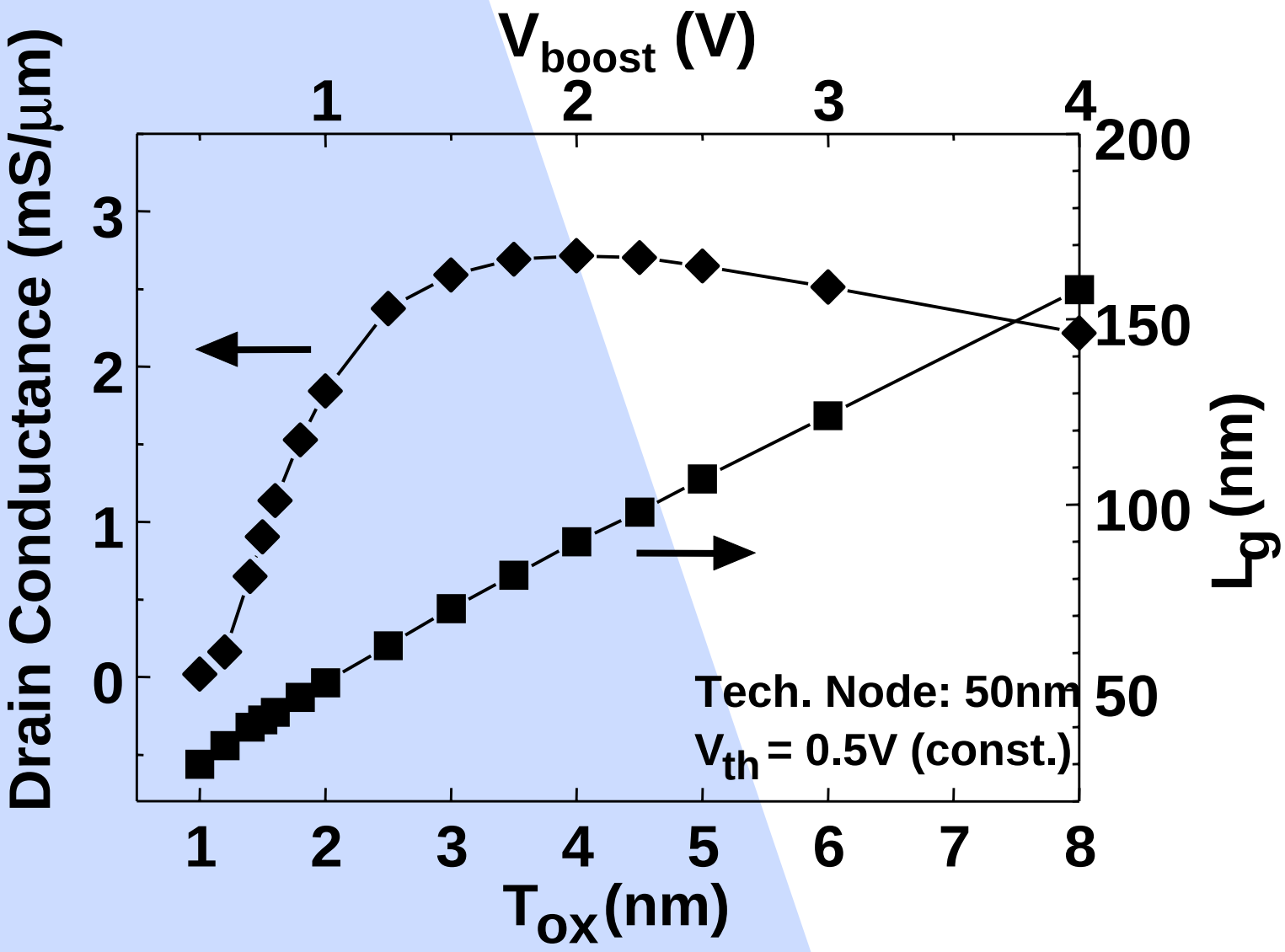
Boosted-Gate MOS (BGMOS)

Device / circuit collaborative approach



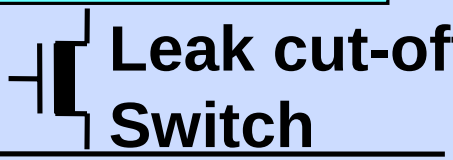
T.Inukai, M.Takamiya, K.Nose, H.Kawaguchi, T.Hiramoto and T. Sakurai, "Boosted Gate MOS (BGMOS): Device/Circuit Cooperation Scheme to Achieve Leakage-Free Giga-Scale Integration," CICC'00, to be published, May 2000.

Leak switch optimization



SRAM doesn't accept MTCMOS & VTCMOS

SRAM
スタンバイ時にメモリ情報が消える

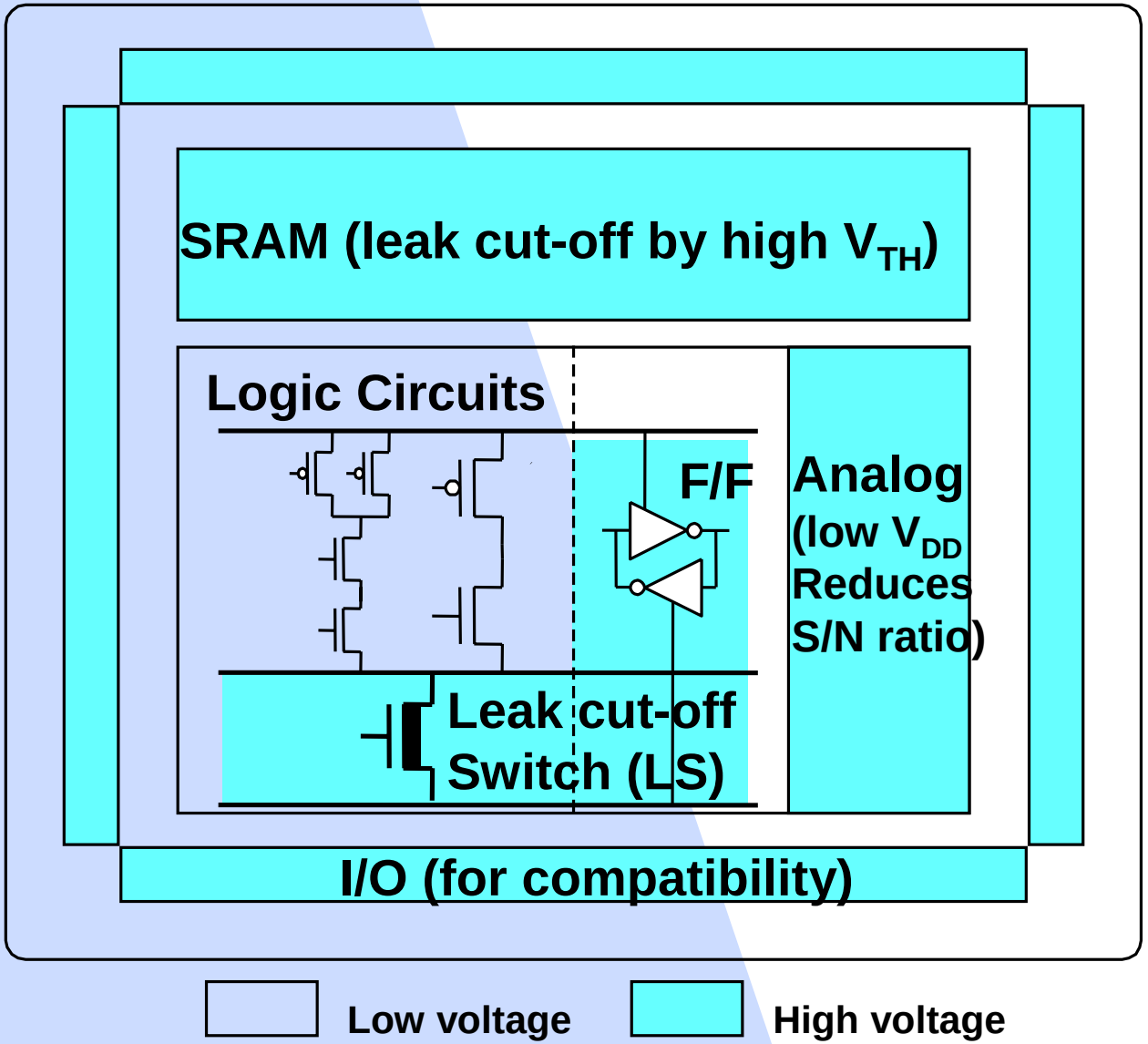


MTCMOS

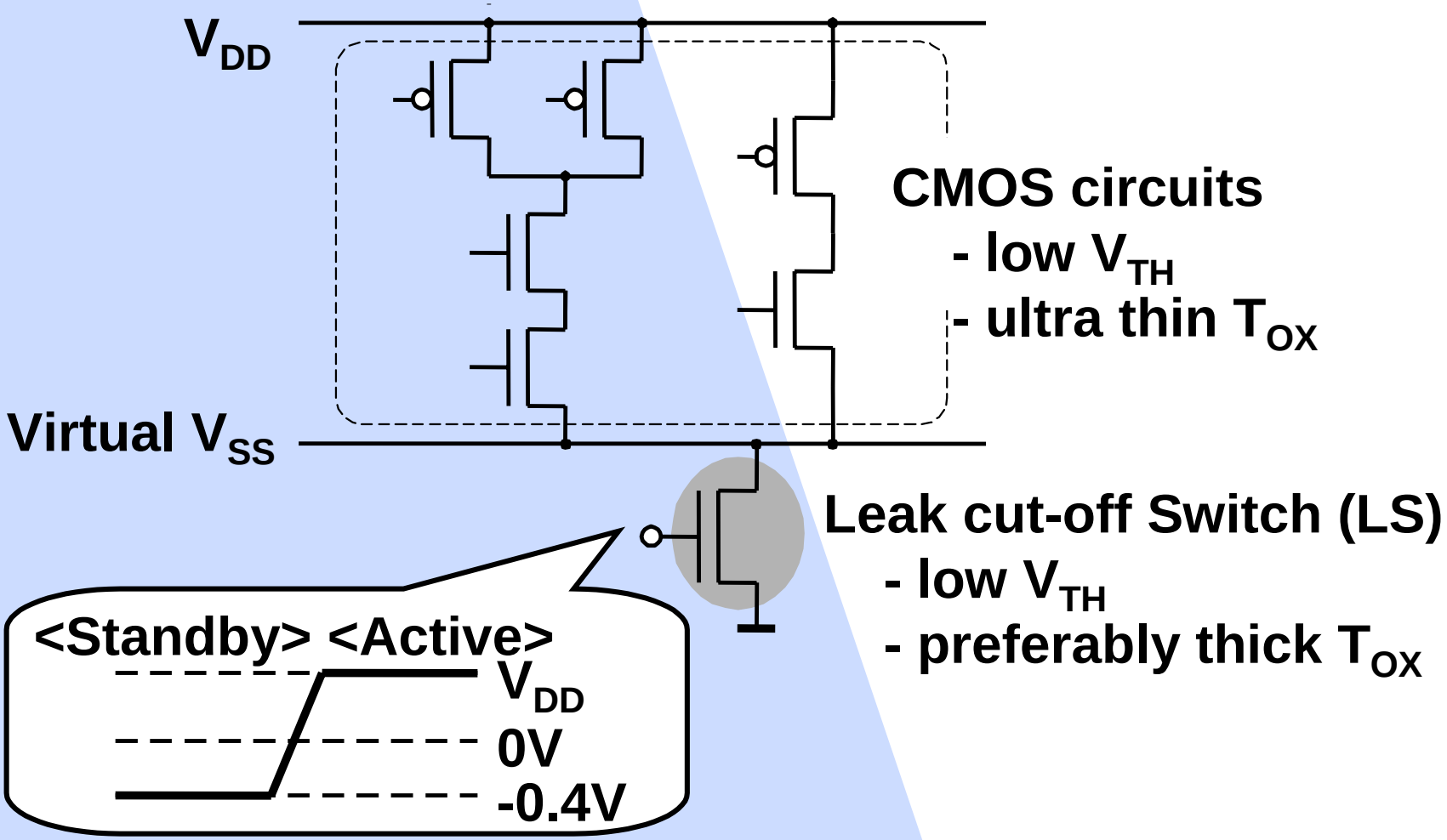
SRAM
アクティブ時にリーク電流によって多大な電力を消費

VTCMOS

Dual oxide thickness

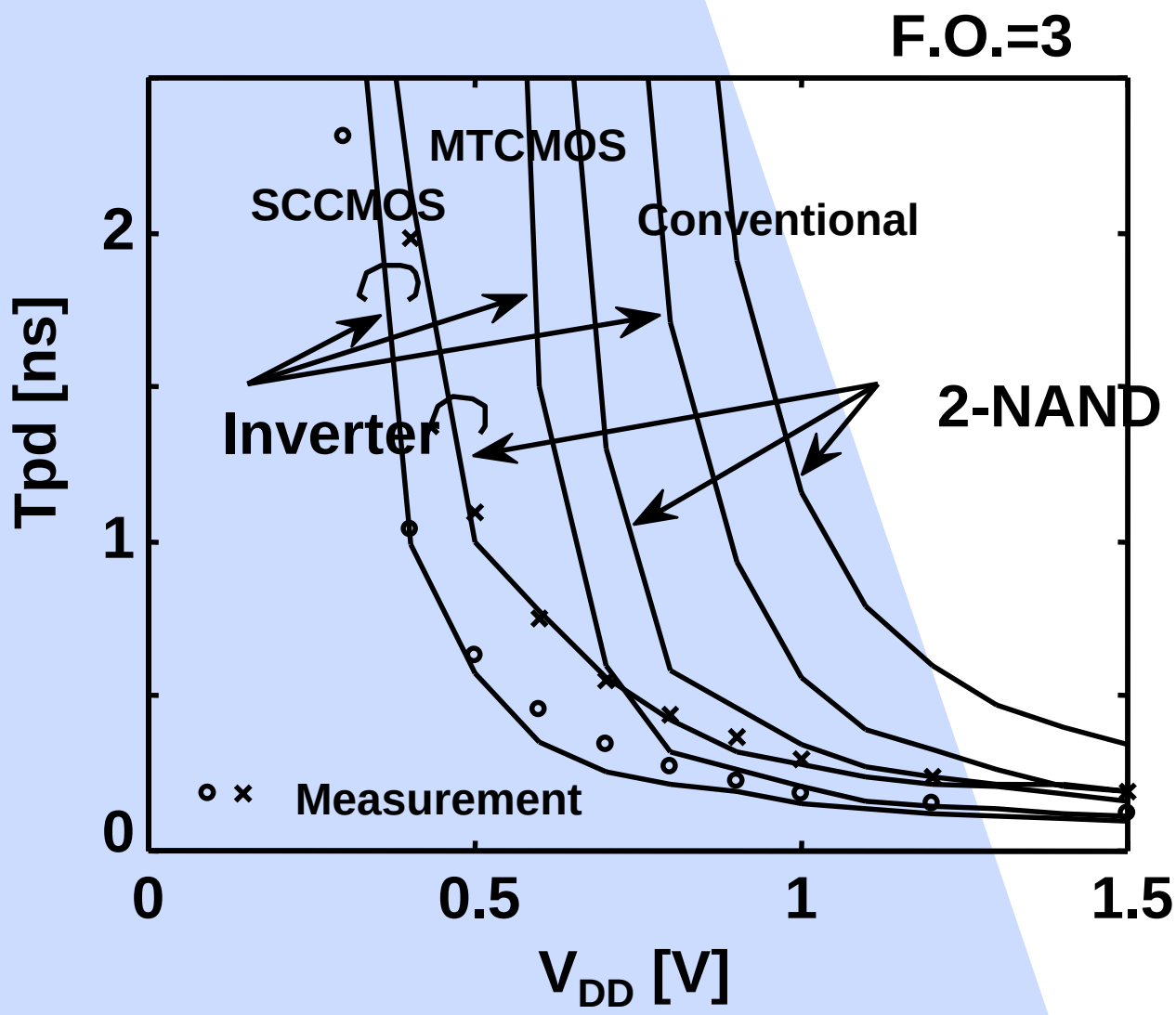


Super Cut-off CMOS (SCCMOS)



H.Kawaguchi and K.Nose, T.Sakurai, "A CMOS Scheme for 0.5V Supply Voltage with pico-Ampere Standby Current," 1998 ISSCC, Digest of Tech. Papers, pp.192-193, Feb. 1998.

Delay characteristics (inverter & NAND)

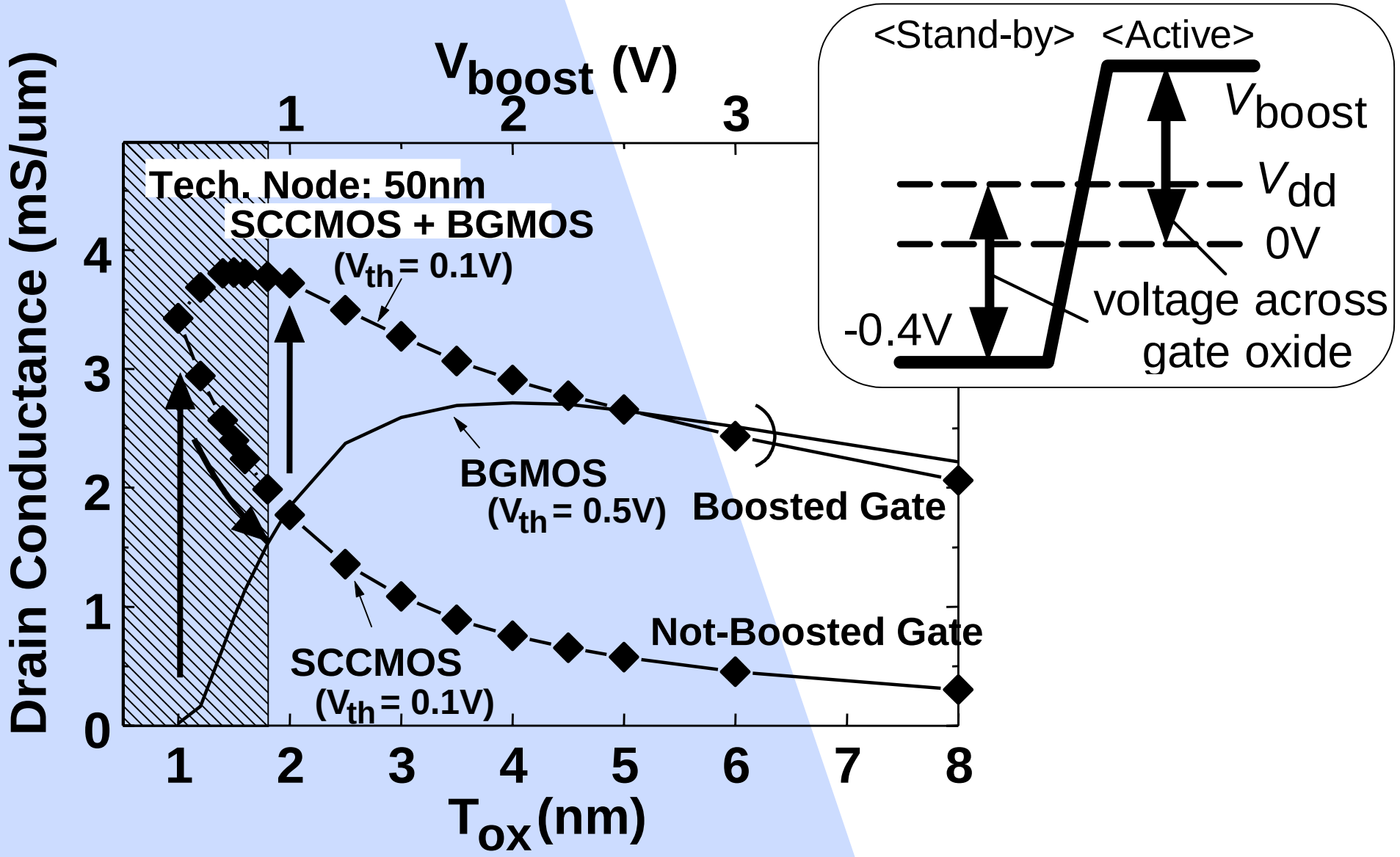


SCCMOS
0.2V V_{TH} circuit
with 0.2V V_{TH}
cut-off MOSFET

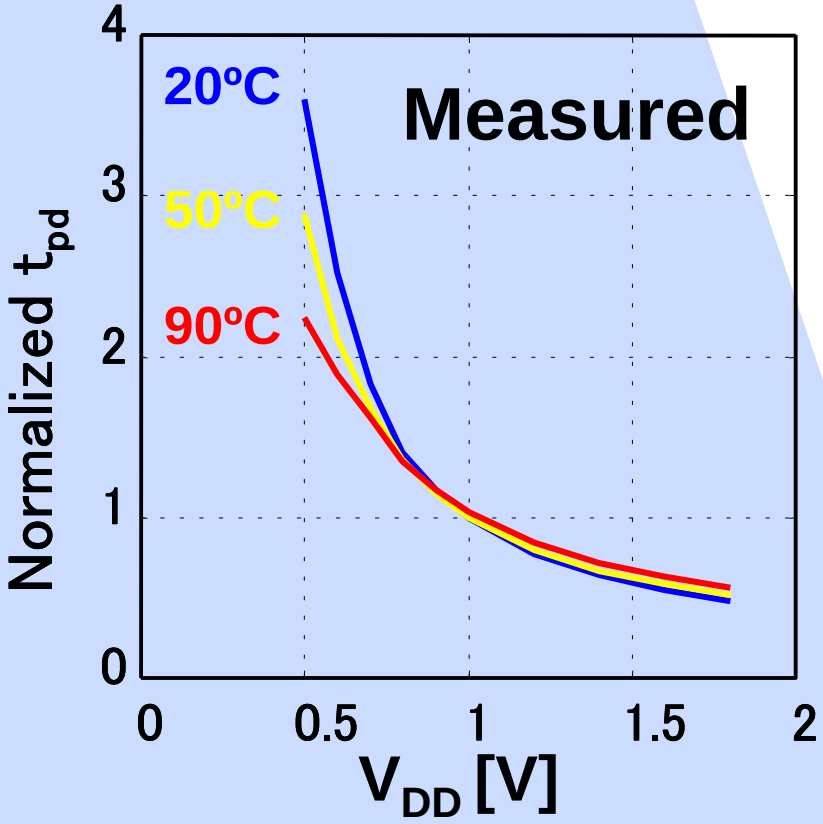
MTCMOS
0.2V V_{TH} circuit
with 0.6V V_{TH}
cut-off MOSFET

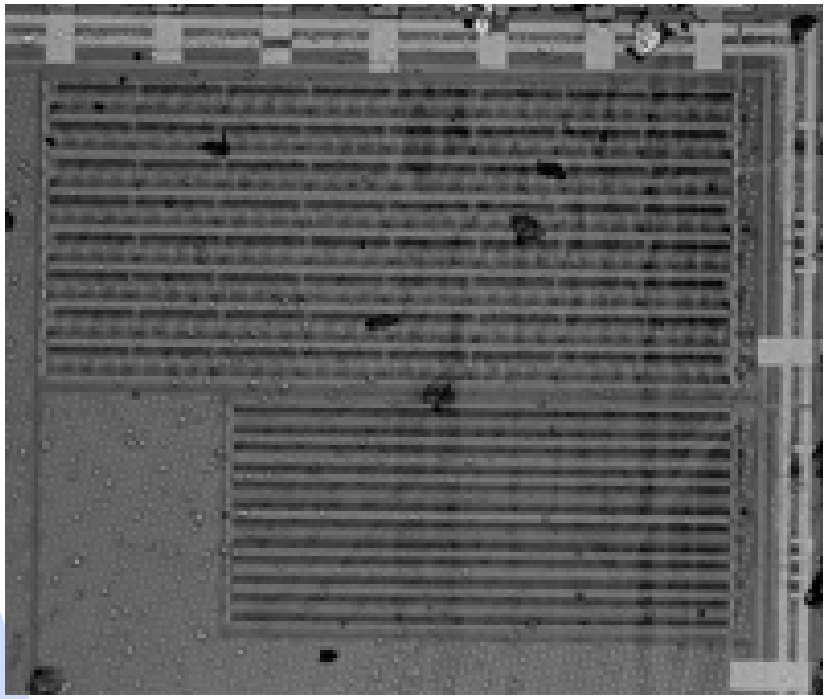
Conventional
All 0.6V circuit
No cut-off
MOSFET

SCCMOS + BGMOS



Measurement of 32bit full adder





Photograph of 32bit FA

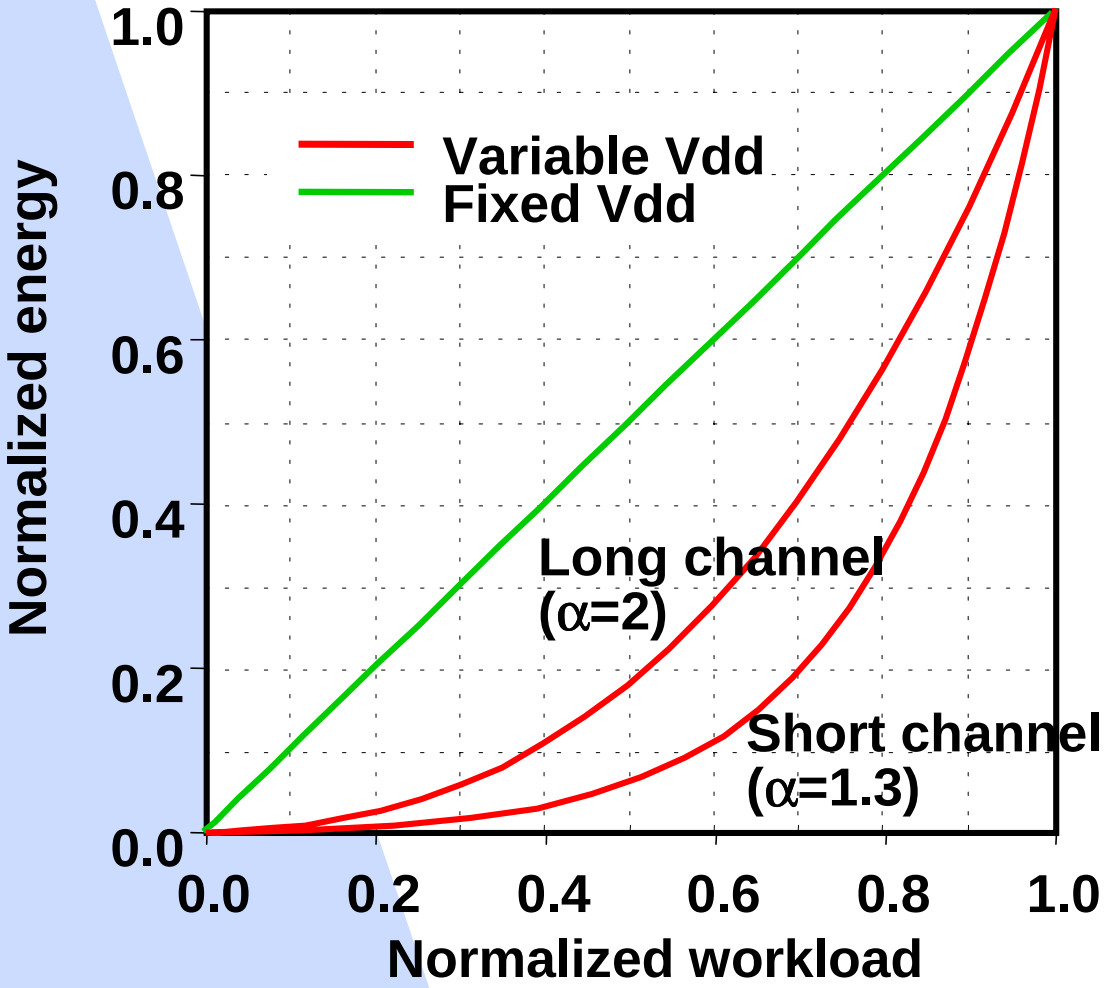
0.3μm CMOS

K.Kanda, K.Nose, H.Kawaguchi, and T.Sakurai,"Design Impact of Positive Temperature Dependence of Drain Current in Sub 1V CMOS VLSI's",CICC99, pp.563-566, May 1999.

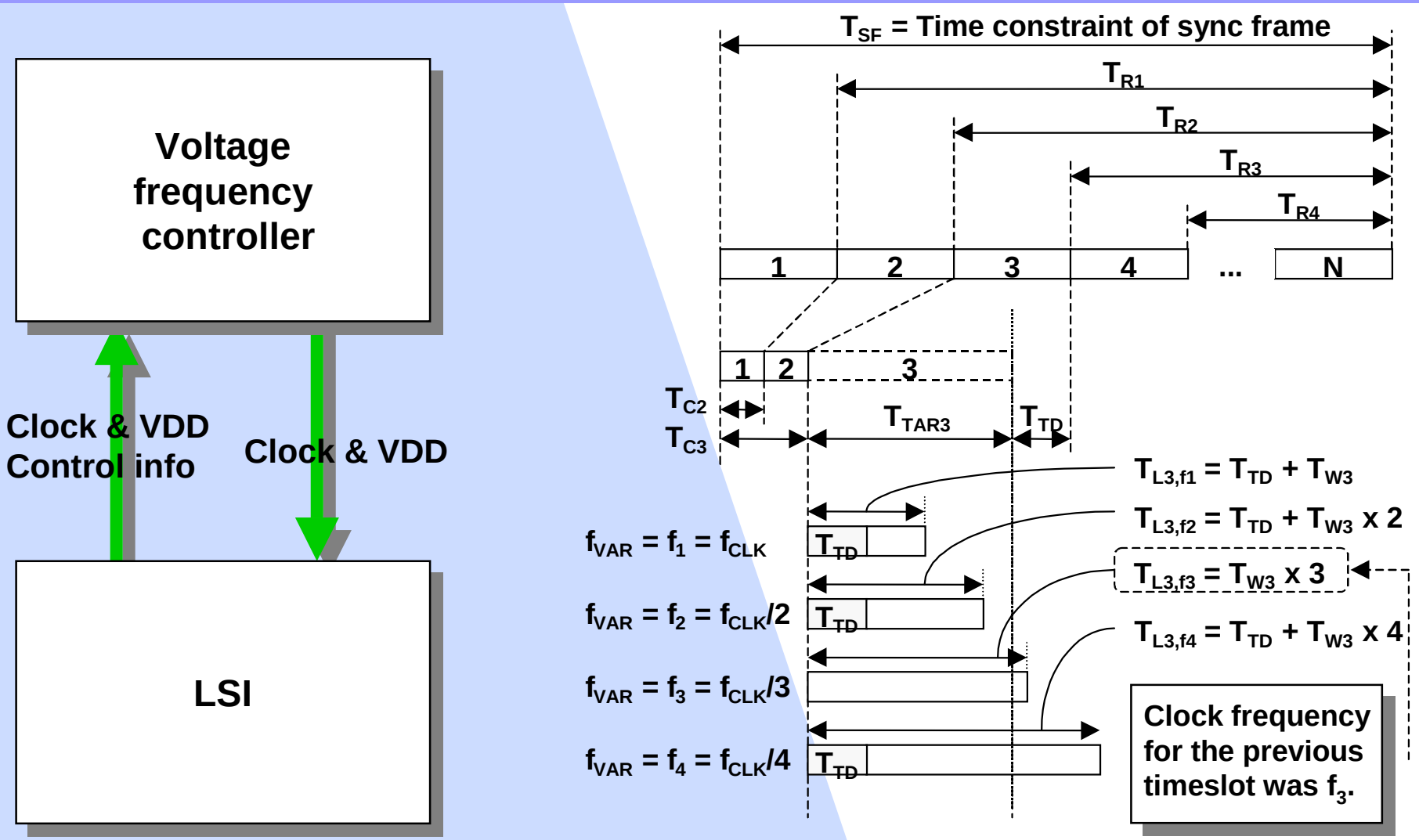
If you don't need to hussle, V_{DD} should be as low as possible

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.



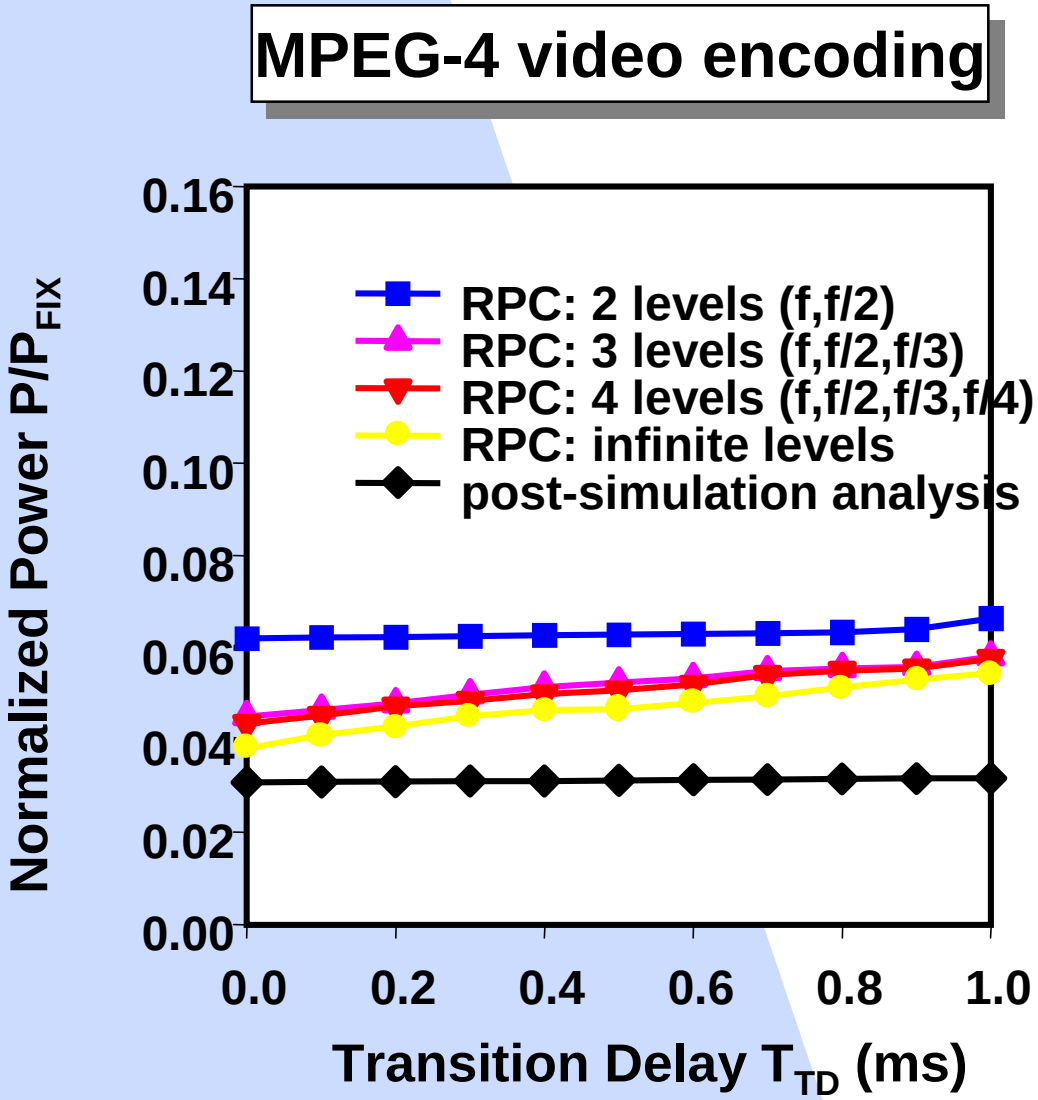
Voltage hopping for low-power



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.

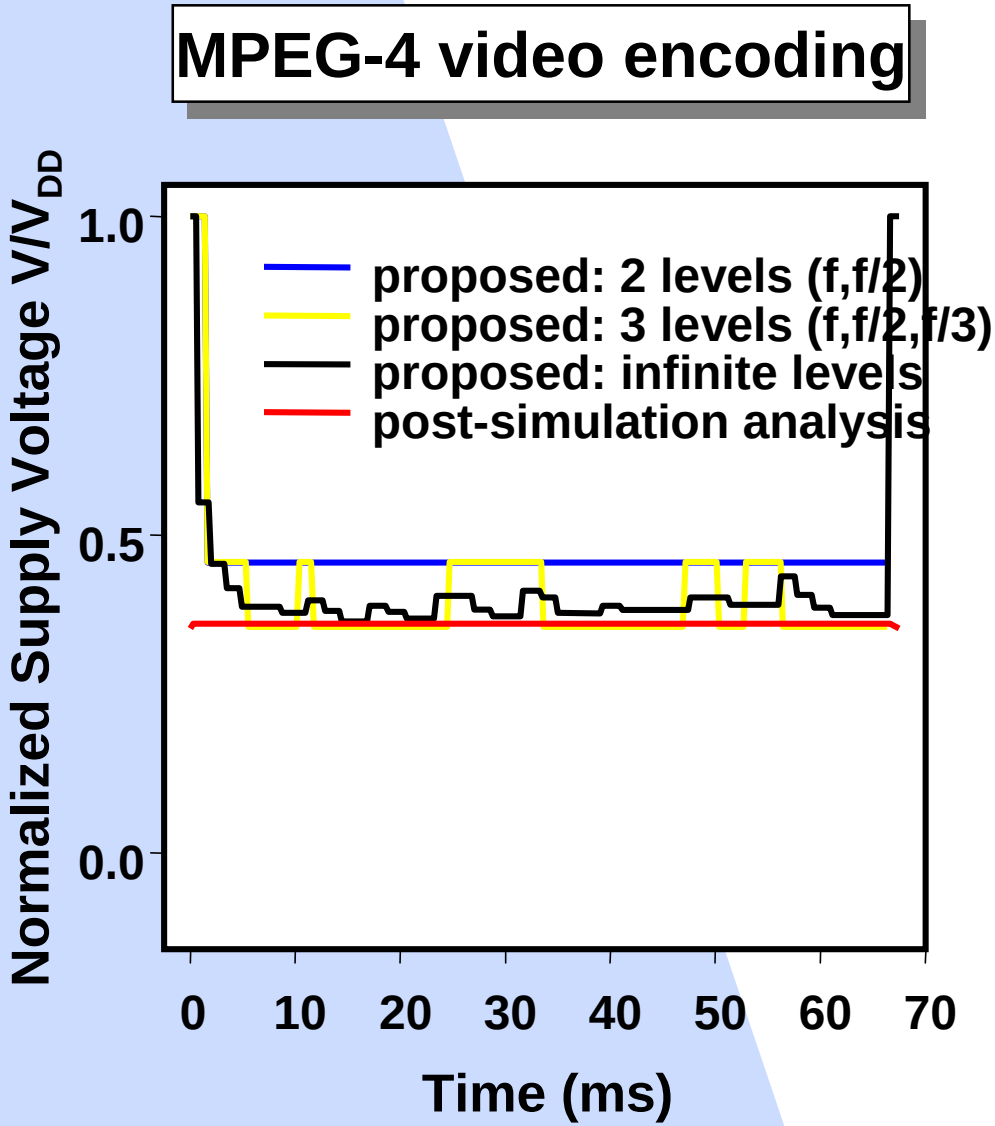
S.Lee and T.Sakurai, “Run-time Voltage Hopping for Low-power Real-time Systems,” to be published, DAC'00, June 2000.

Transient voltage waveform



Simulation results

Transient voltage waveform

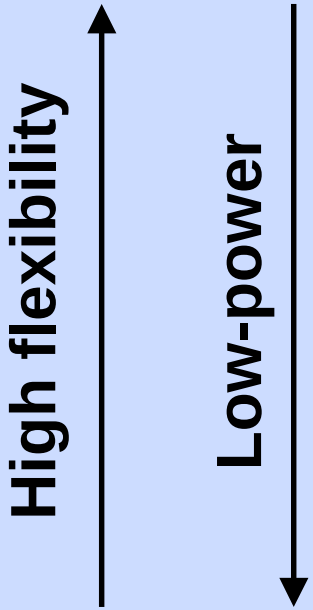


Simulation results

1 sync frame = 33 timeslots = 66.66667 ms

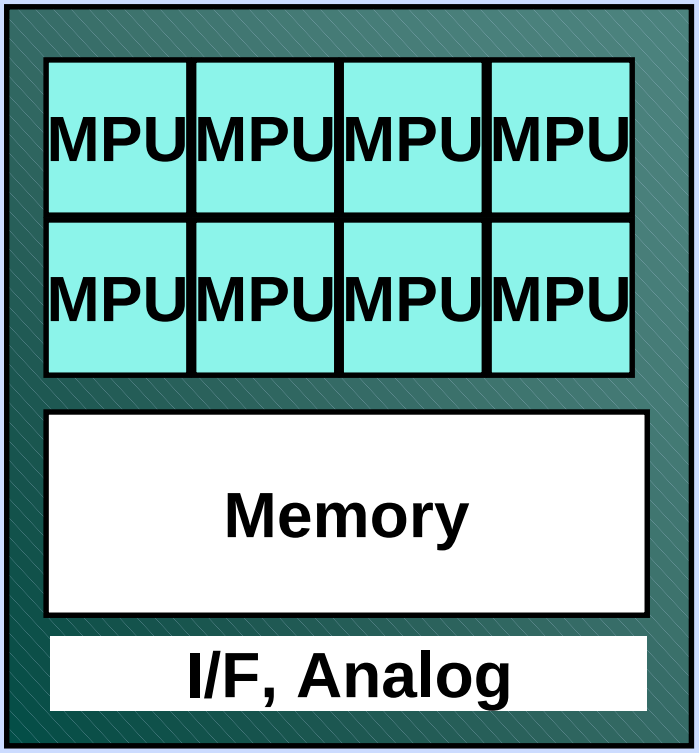
Approach to low-power LSI

Example of MPEG2 decoding



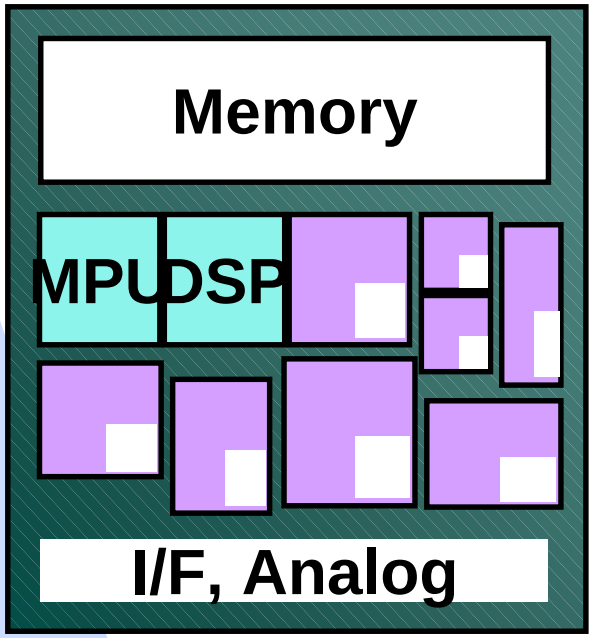
- Processor (software)
  ～25W
- DSP
  ～4W
- Dedicated sytem LSI (SW/HW)
  ～0.7W

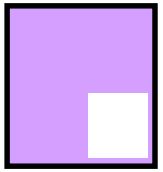
Homogeneous vs. Heterogeneous



Homogeneous Architecture

(High flexibility)



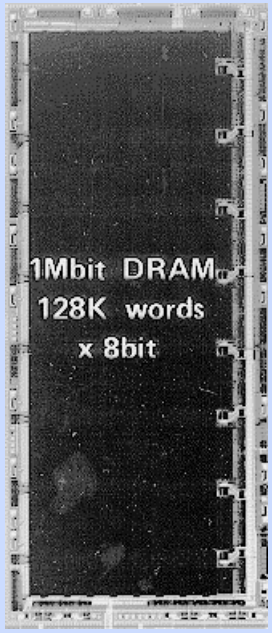


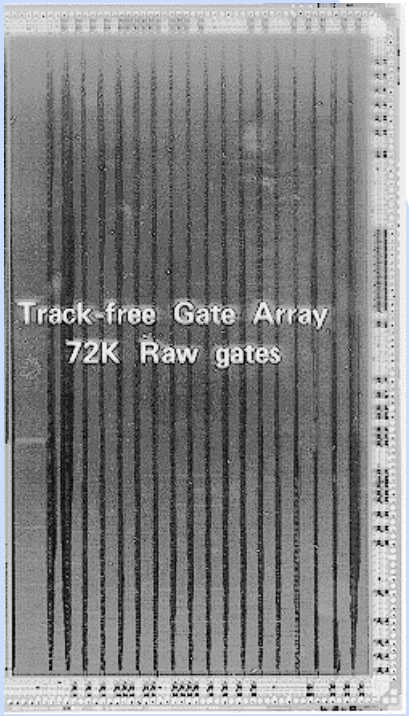
Special Engine

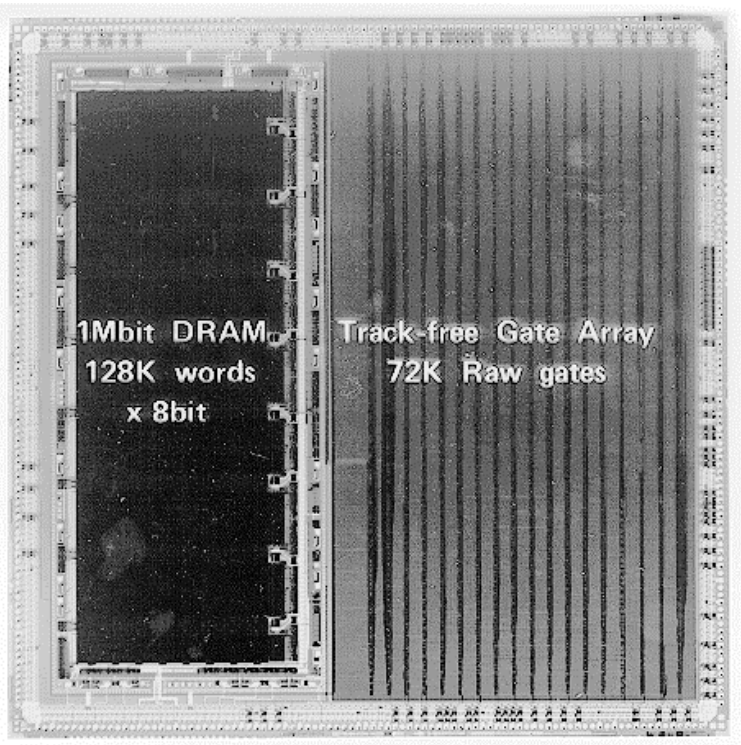
Heterogeneous Architecture (System LSI)

(Low-power, more efficient)

DRAM Embedding





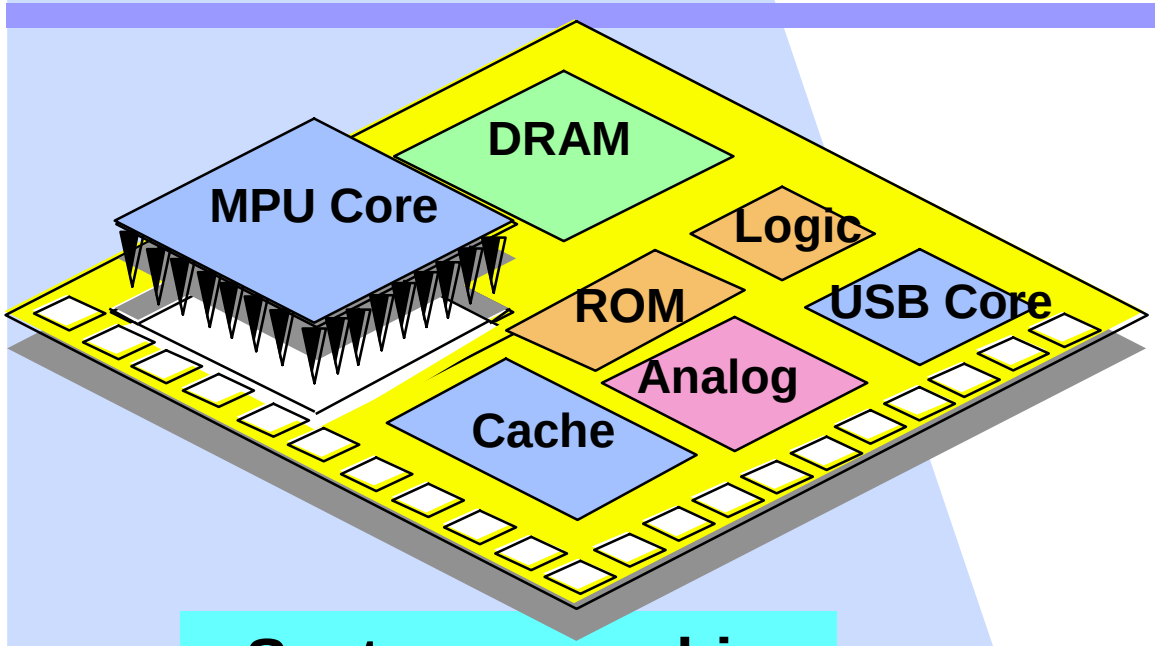


DRAM **Processor** **System LSI**

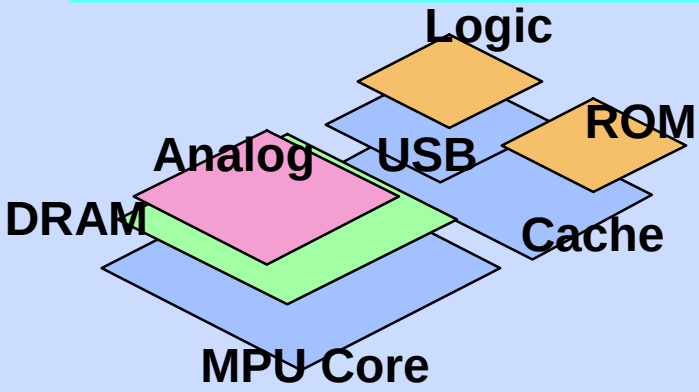
K.Sawada, T.Sakurai, et al, "A 72K CMOS Channelless Gate Array with Embedded 1Mbit Dynamic RAM," in Proc. CICC'88, pp.20.3.1-20.3.4, May 1988.

- **Two orders of magnitude improvement in bandwidth and power**

3-Dimensional assembly



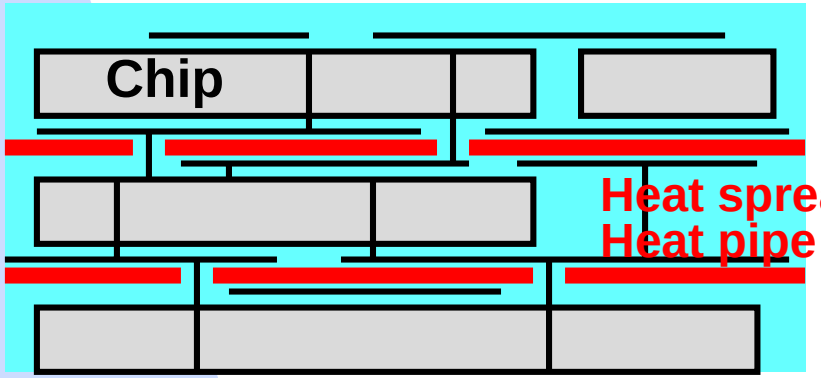
System on a chip



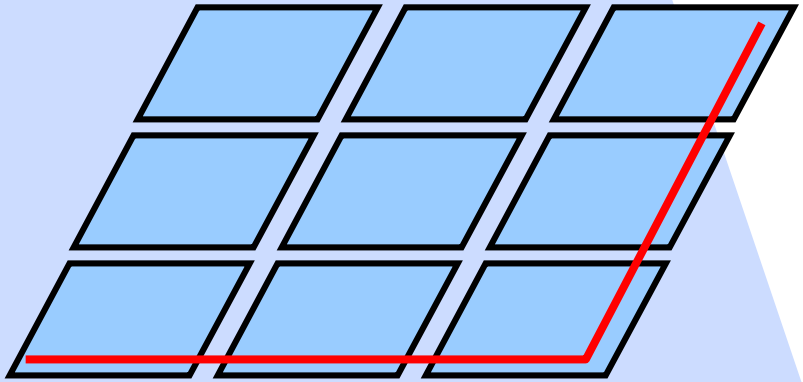
3-D assembly

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

- Heat dissipation is an issue



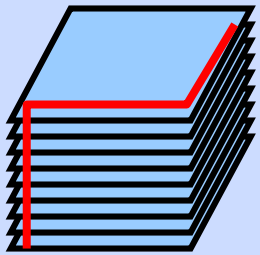
Shorter interconnect in 3-D assembly



System on a chip

$$\frac{\#\textit{of devices in } d\ (3D)}{\#\textit{of devices in } d\ (2D)} = \frac{1}{3}\left(2\frac{d}{h}+\frac{h}{d}\right)$$

$$\approx \frac{2}{3}\left(\#\textit{of stacked chips in } d\right)$$



3-D assembly

d: Manhattan distance

h: Height between chips

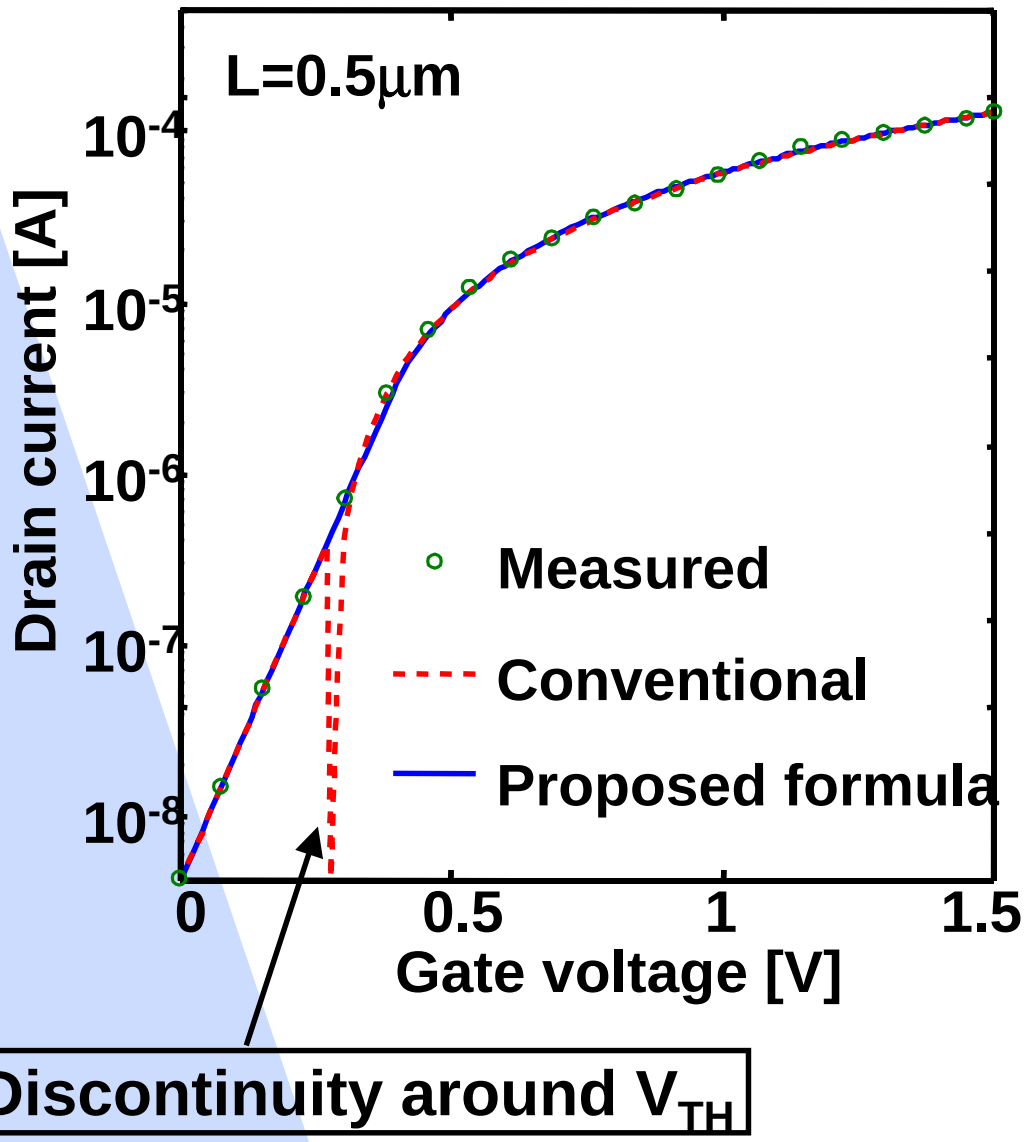
Drain current model

$$I_D = I_0 e^{\alpha\left(\frac{V_{GS}-V_{TH}}{\alpha N_S}\right)^{\alpha}}$$

$$(V_{GS} \geq V_{TH} + \alpha N_S)$$

$$I_D = I_0 e^{\frac{V_{GS}-V_{TH}}{N_S}}$$

$$(V_{GS} \leq V_{TH} + \alpha N_S)$$



Optimum V_{DD} and V_{TH}

The optimum solution is obtained by $\frac{\partial P_{max}}{\partial V_{DD}} = 0$ **.**

$$V_{THopt} = -N_S \ln\left(\frac{2afC_L N_S}{I_0}\frac{\alpha}{\alpha-\chi}\right) \qquad \chi = \left(\frac{fL_d C_L K}{\beta}\right)^{1/\alpha}$$

(Optimum $V_{TH,min}$)

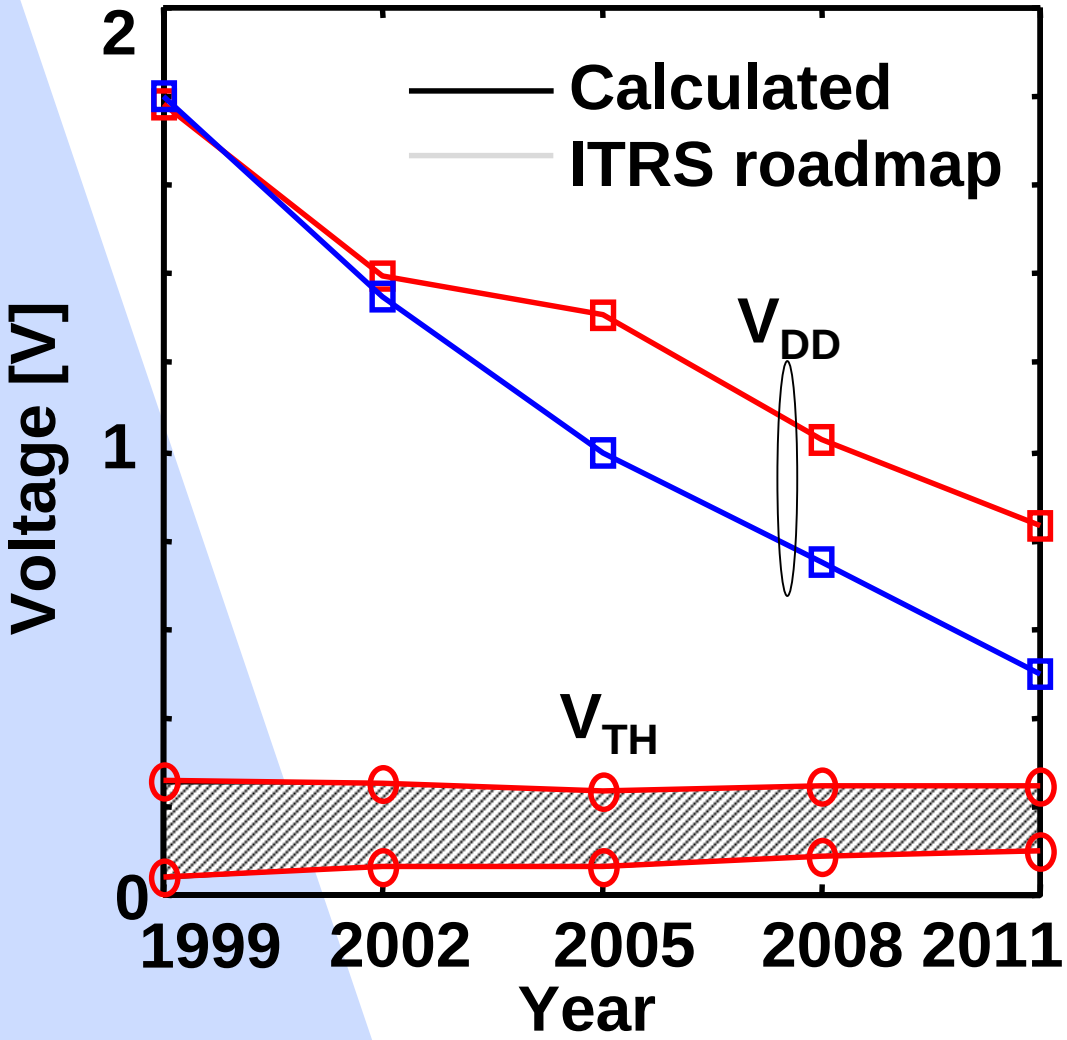
$$V_{DDopt} = \frac{-N_S \ln\left(\frac{2afC_L N_S}{I_0}\frac{\alpha}{\alpha-\chi}\right) + \Delta V_{TH} + \kappa\Delta T + \frac{\alpha-1}{\alpha}\chi}{1-\frac{\chi}{\alpha}}$$

$$\frac{P_{LEAK,\max}}{P_{\max}} = \frac{P_{LEAK,\max}}{P_D + P_{LEAK,\max}} \cong \frac{2N_S\alpha}{2N_S\alpha+\alpha-1} \cong 0.3 \quad \left(\begin{array}{c}\alpha = 1.3 \\ N_S = 80mV/decade\end{array}\right)$$

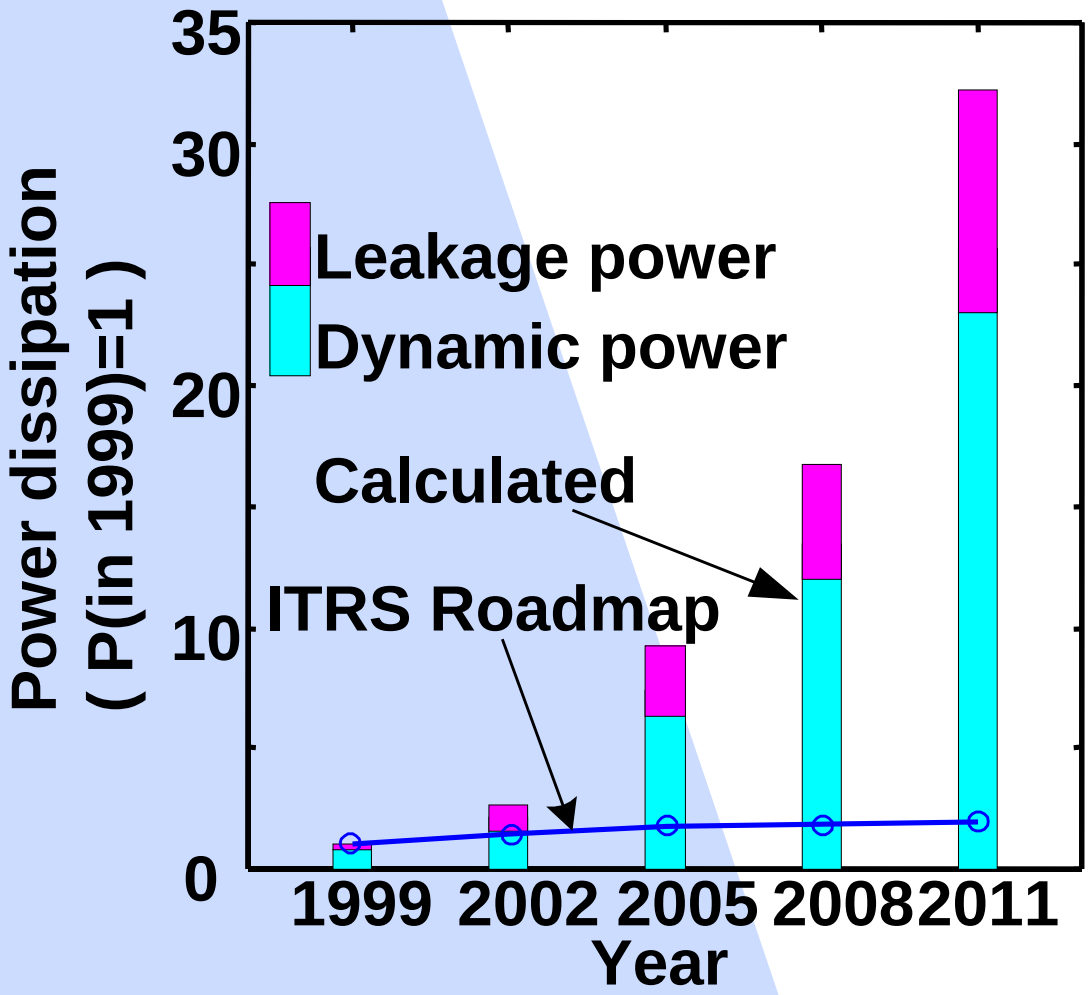
Problems of scaling scenario (no modification scenario)

In order to maintain L_d=20, V_{DD} > 0.8V in 2011.

N_{CHIP} in 2011 is 70 times larger than N_{CHIP} in 1999.

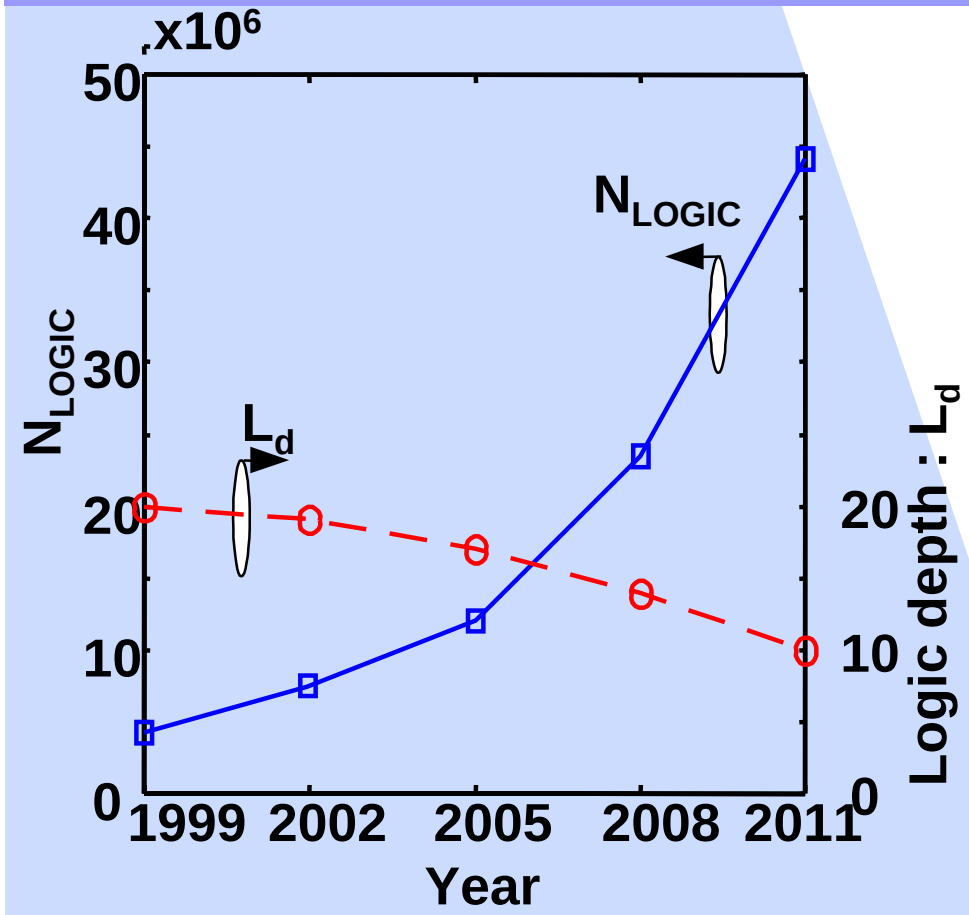


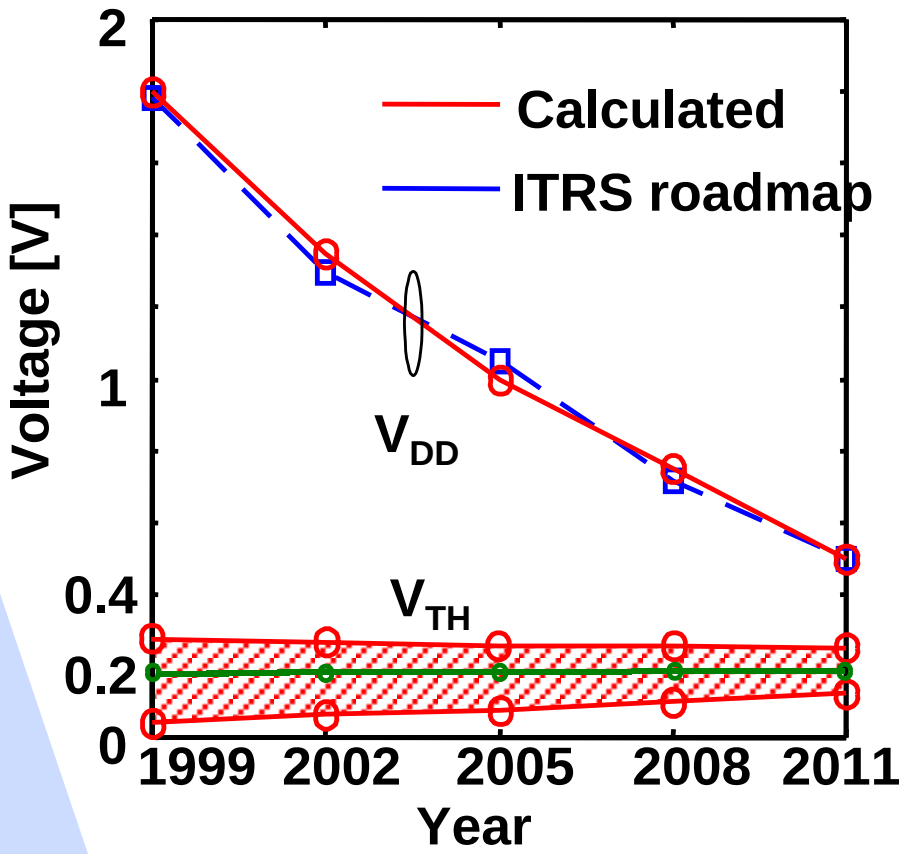
Calculation of future trend



- Power dissipation in 2011 is 32 times as large as that in 1999 when there is no modifications.

Proposed scaling scenario





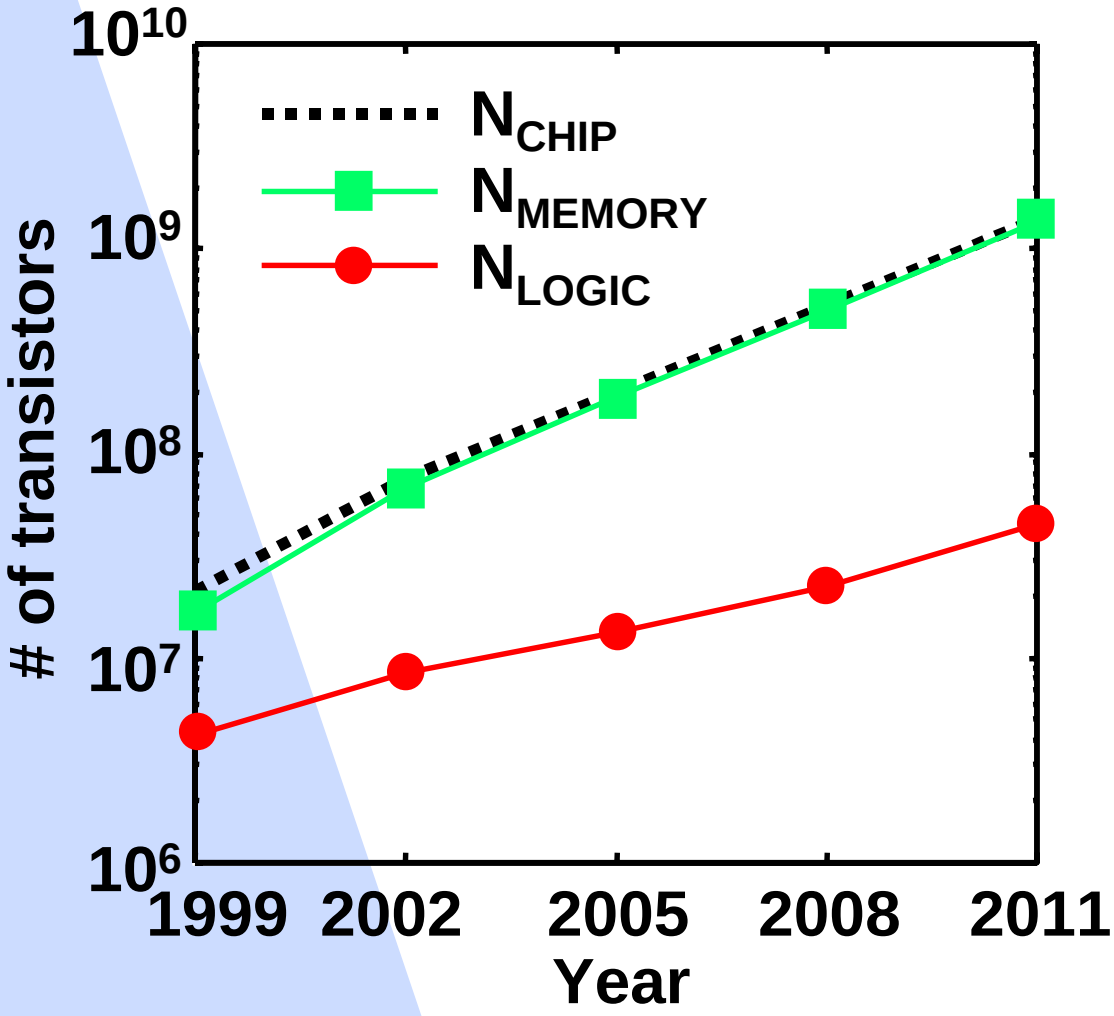
Optimum $V_{TH,min}$ 0.06V (in 1999) → 0.11V (in 2011)

Optimum $V_{TH,max}$ 0.26V (in 1999) → 0.22V (in 2011)

Target V_{TH} 0.21V (in 1999) → 0.20V (in 2011)

Future trend of N_{LOGIC} and N_{MEMORY}

	1999	2011
$\frac{N_{LOGIC}}{N_{CHIP}}$	20%	3%
$\frac{N_{MEMORY}}{N_{CHIP}}$	80%	97%

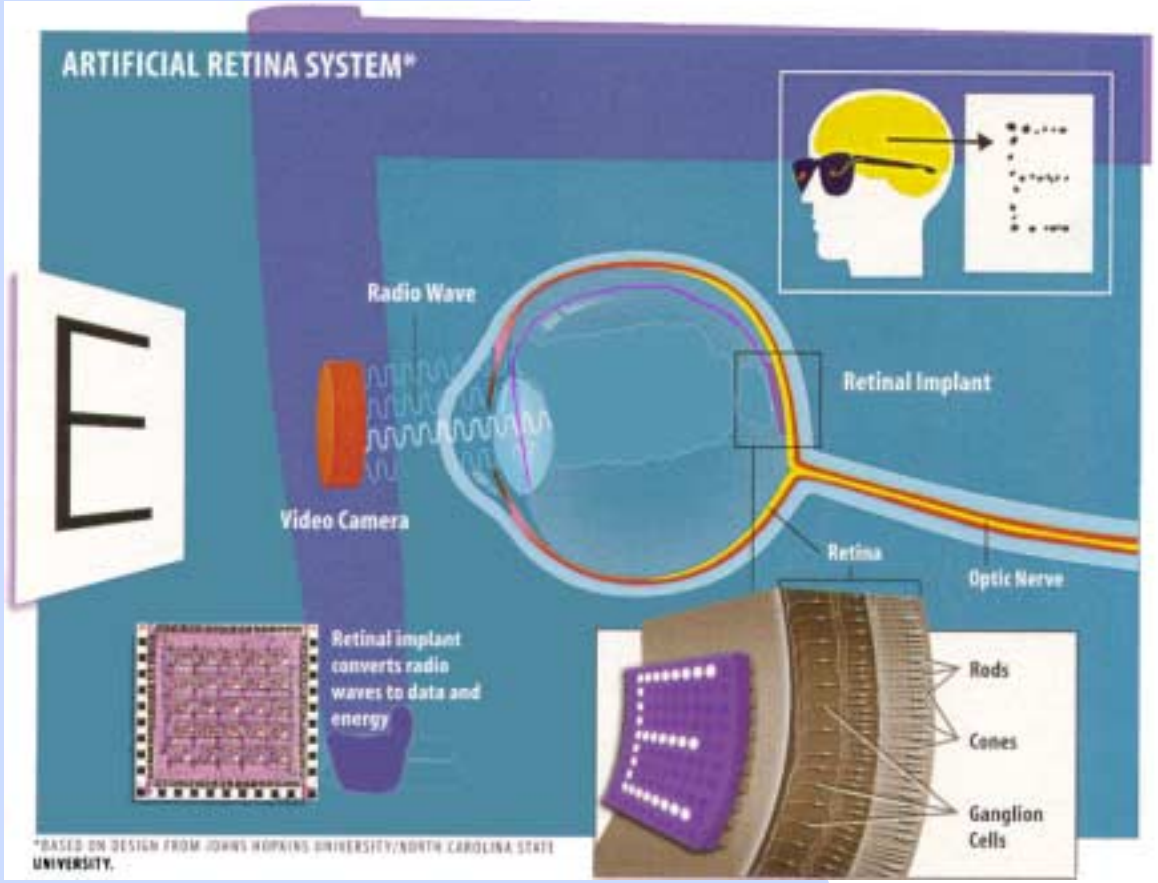


Summary

- Device / circuit collaborative approach: Dual-oxide Boosted Gate MOS (BGMOS)
- Circuit / software collaborative approach: Voltage hopping
- System LSI approach is effective for low-power
- Memory-rich architectures are to be sought.

Prosthesis - Dual Intraocular Units

NC STATE UNIVERSITY



Retinal Prosthesis

Reference for low-power design & System LSI

Low-power high-speed LSI design & technology

「低消費電力、高速LSI技術」

Realize publishing company,

ISBN4-89808-004-9 C3055 ¥56000E

Phone: +81-3-3815-8511, Fax: +81-3-3815-8529

System LSI – Applications and Technology

「システムLSIーアプリケーションと技術」

Science Forum publishing,

ISBN4-916164-30-X C3000 ¥48000E

Phone: +81-3-5689-5611, Fax: +81-3-5689-5622