

# Low Power Techniques

**Author : Dr. Koichiro Ishibashi**

**Updated by : Hoang Trinh**

**Reviewed by: Dr. Hai Pham**

**Renesas Design Vietnam Co., Ltd.**

**Backend 3 Group**

For training class on Aug 08, 2018

# Target of this course

- Introduce some popular techniques in Low power design.
- In each technique, we will go through the **basic concept, the benefits and issues.**

# Contents

- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power

# Contents

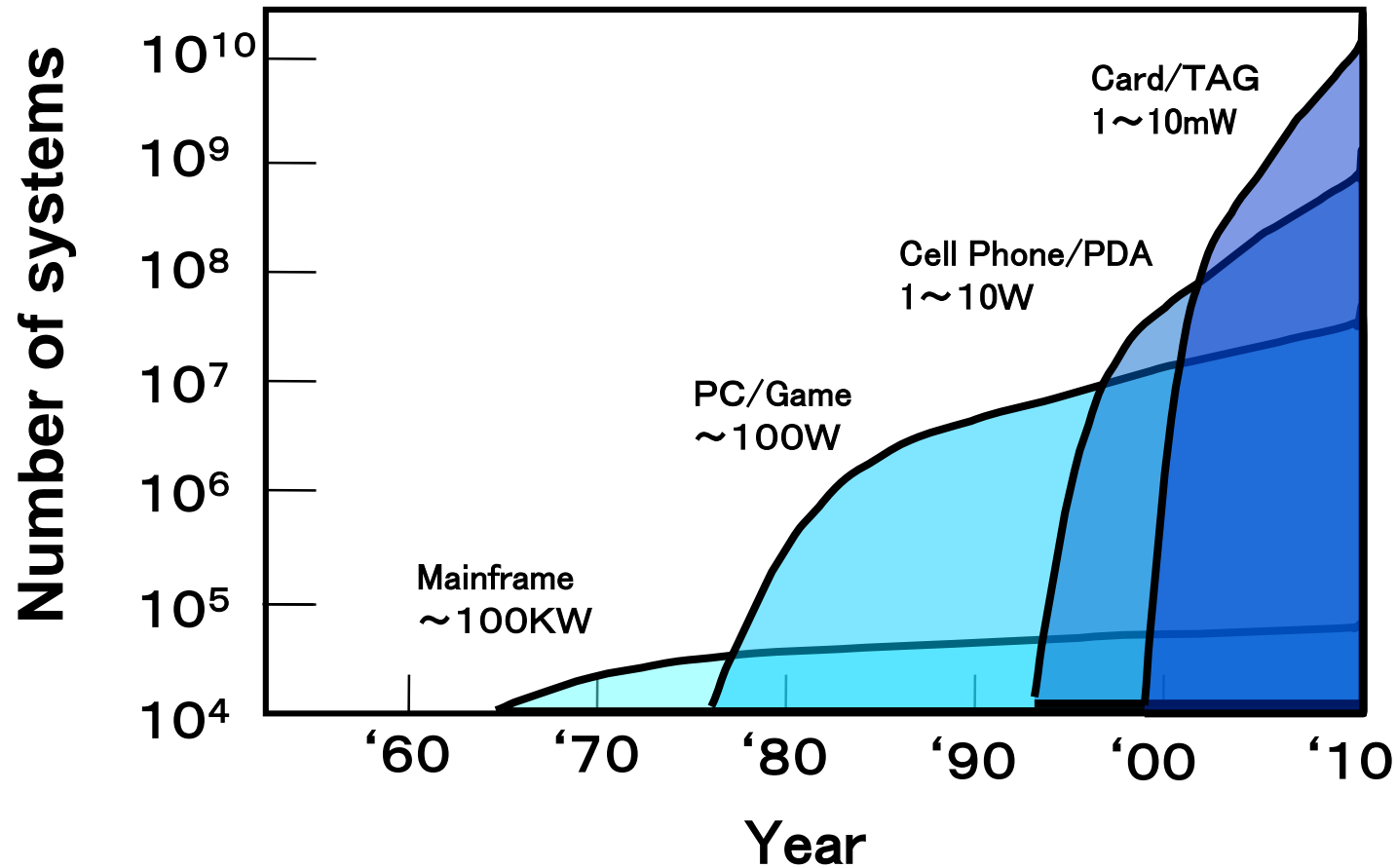
- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power

# Importance of Low Power

Low power has become an important issue in LSI design, due to:

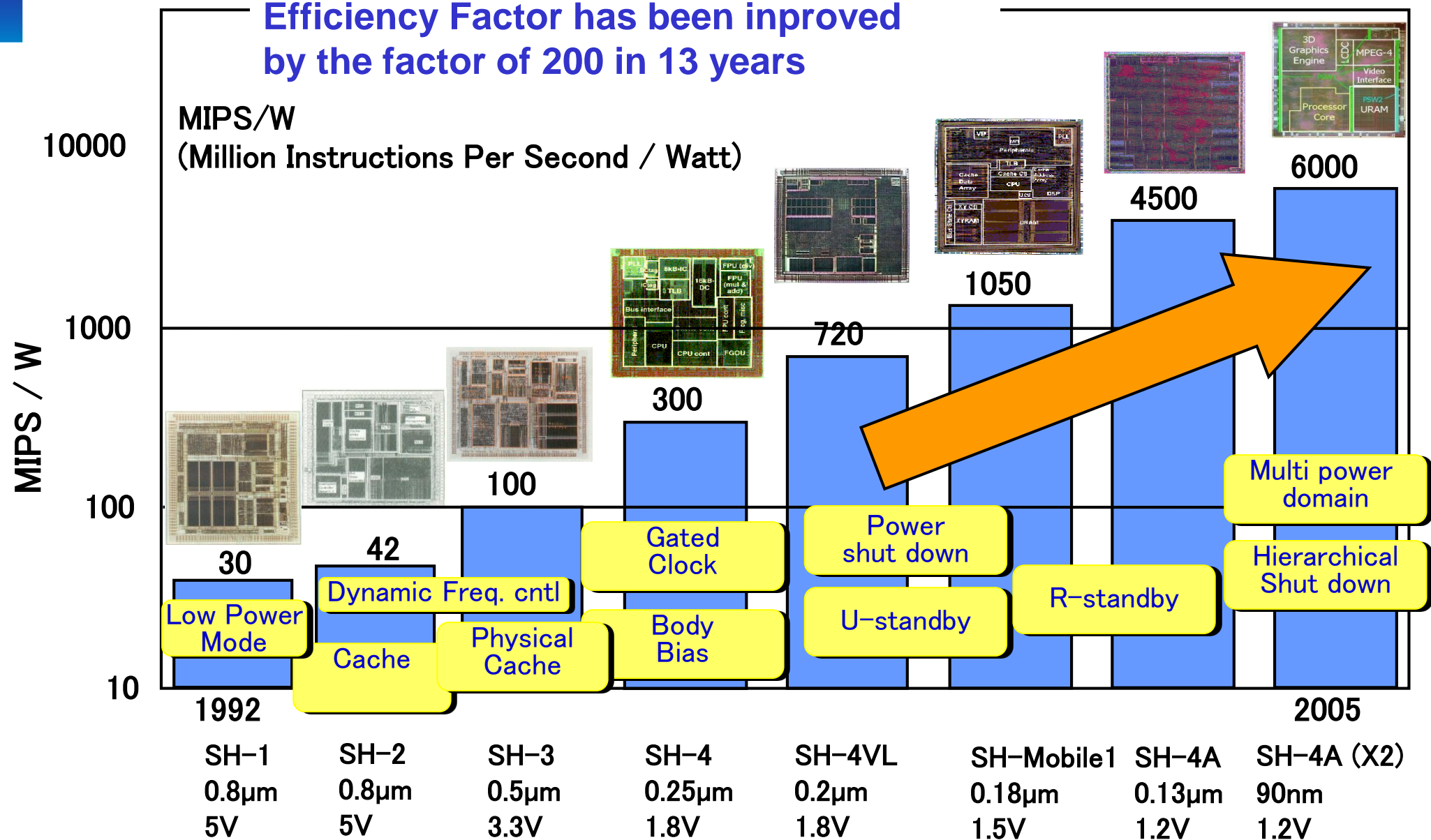
- ❑ **The number of Transistor is getting double every 18 months, leading to the increasing power consumption.**
- ❑ **The operation speed of LSI is higher and higher. The power dissipation is proportional to the clock frequency.**
- ❑ **Power dissipation in form of heat. Lowering power consumption will help to reduce cost for cooling systems.**
- ❑ **Environmental issues have come to be recognized worldwide as extremely important. Every nation, society, and company must collaborate actively to solve these issues.**

# Semiconductor Application Transition



# Power Efficiency of Super H

Efficiency Factor has been improved  
by the factor of 200 in 13 years

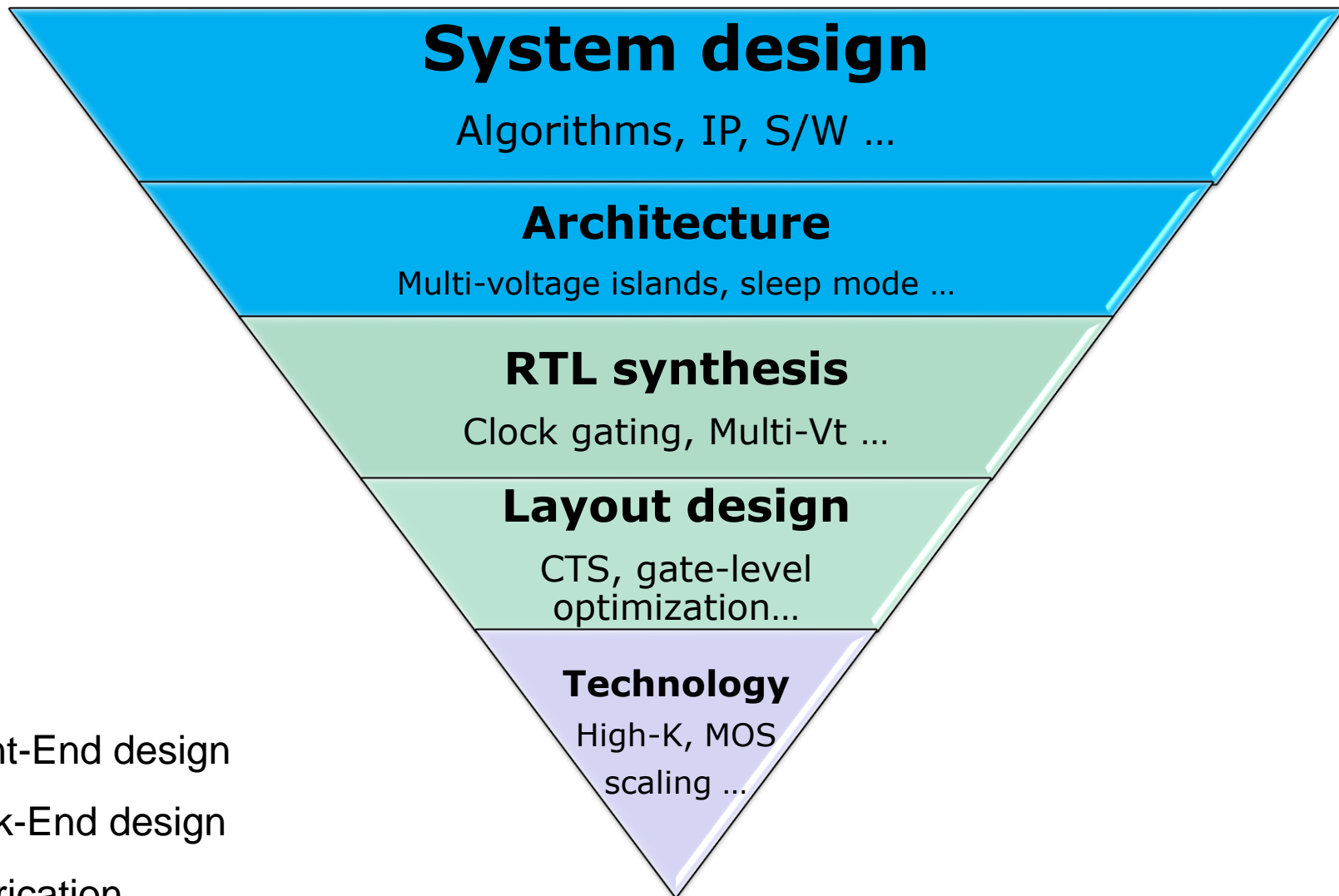


# Contents

- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power



# Where to apply Low Power?



# Power Consumption Calculation

$$P = A N_t C_L V^2 f + N_t I_L V$$

## Dynamic power

- ❖ During the switching of transistors
- ❖ **Depends** on the **clock frequency and switching activity**

## Static power

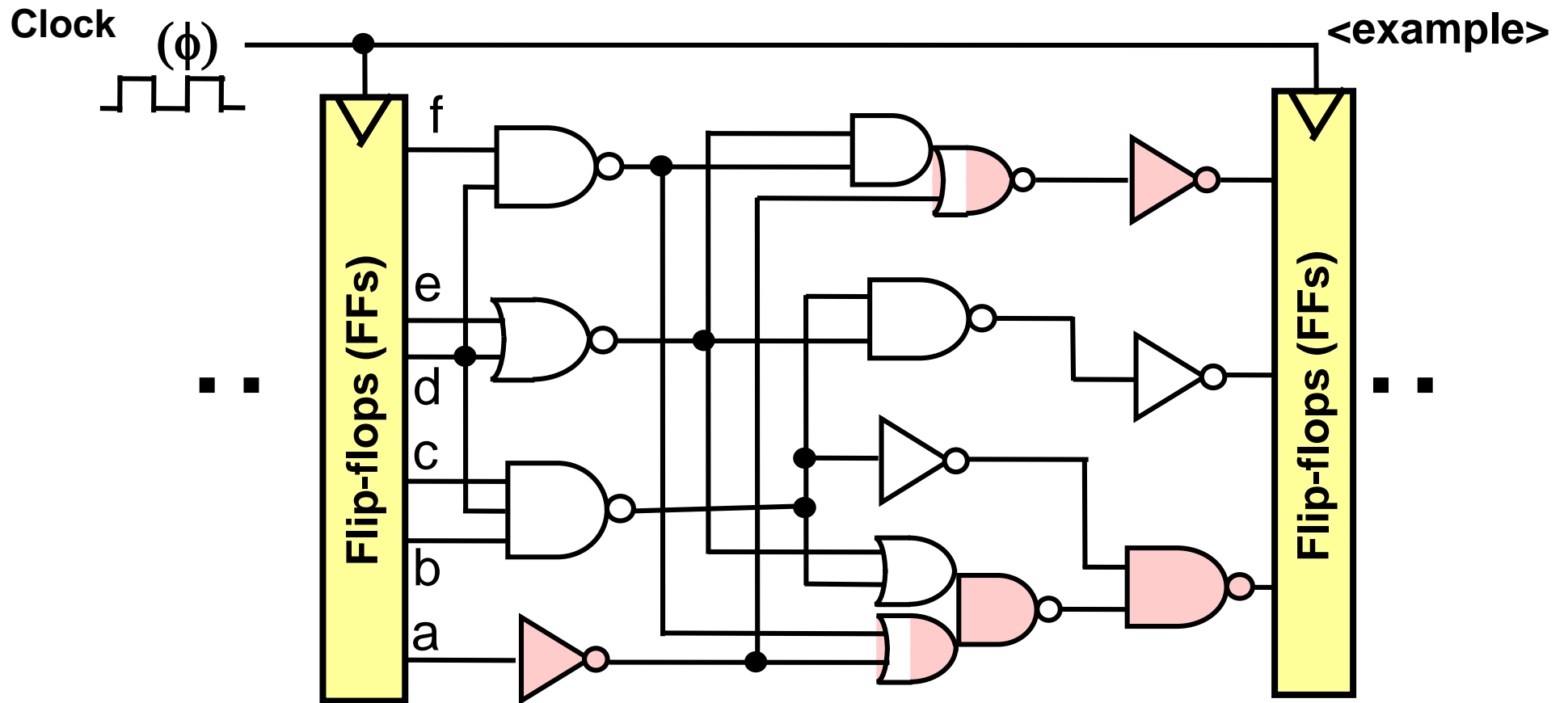
- ❖ Transistor leakage current that flows whenever power is applied to the device
- ❖ **Independent** on the **clock frequency or switching activity.**

where  $I_L$ : Leakage current  
 $A$ : Activity ratio

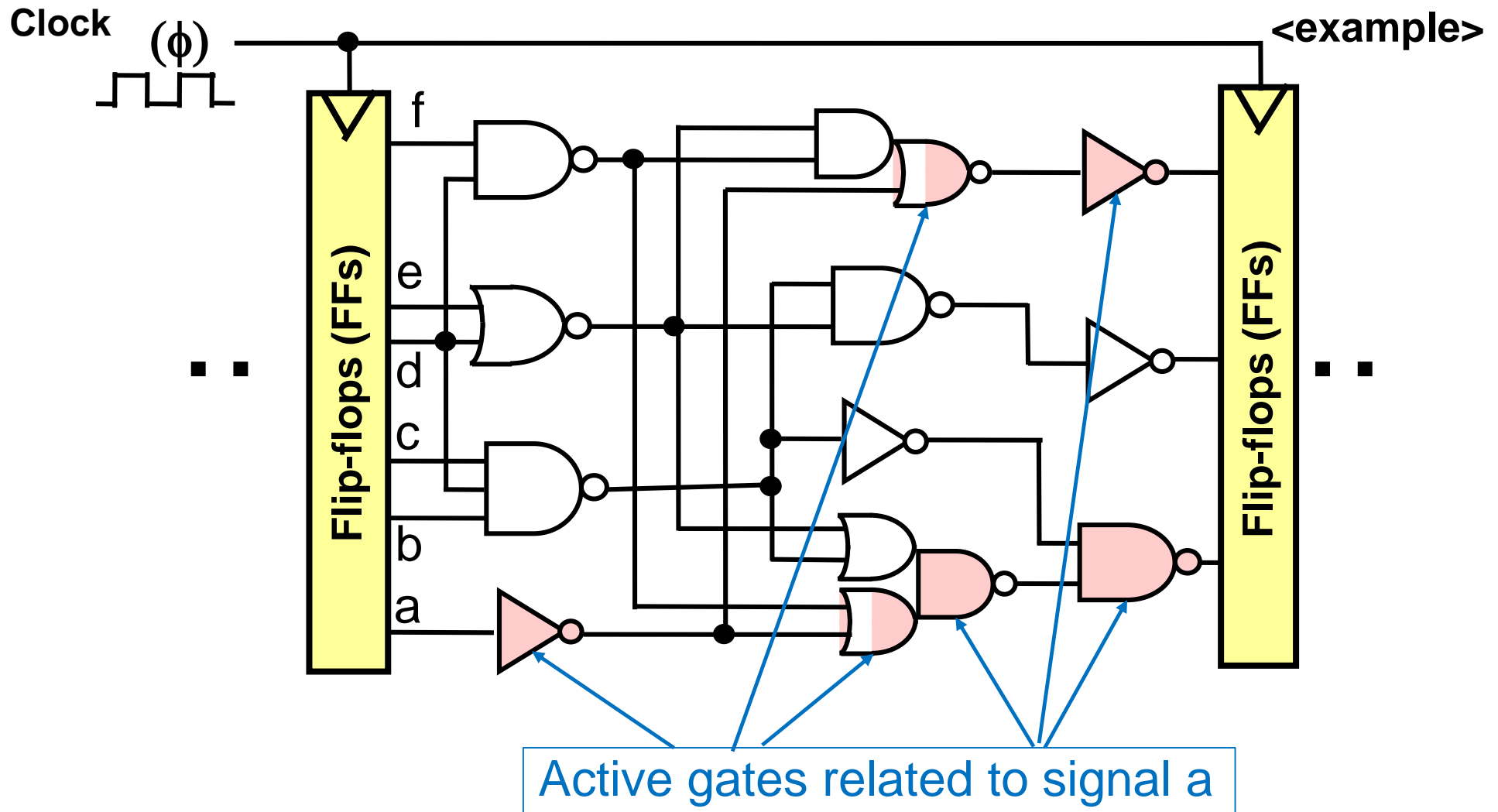
$C_L$ : Average capacitance  
 $N_t$ : Number of total gate

$V$ : supply voltage  
 $f$ : clock frequency

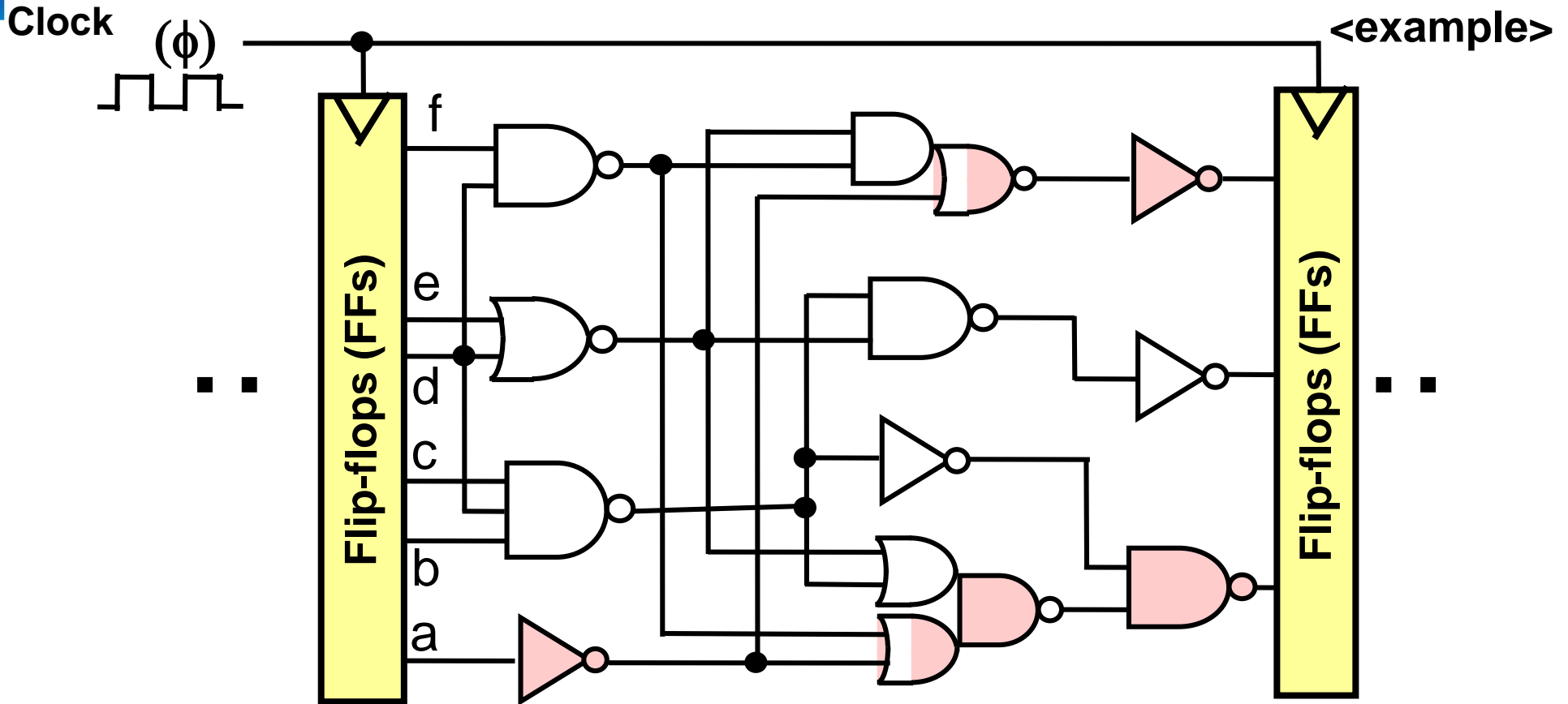
# Explanation - Activity Ratio



# Explanation - Activity Ratio



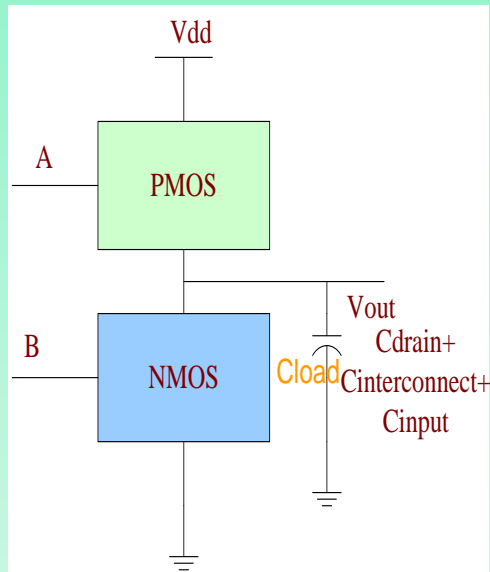
# Explanation - Activity Ratio



**Activity Ratio :  $A = \# \text{ of active gates} / \# \text{ of all gates}$**

# Power Consumption

$$\text{Dynamic power ( } P_{\text{dynamic}} = A N_t C_L V^2 f \text{ )}$$



CMOS inverter

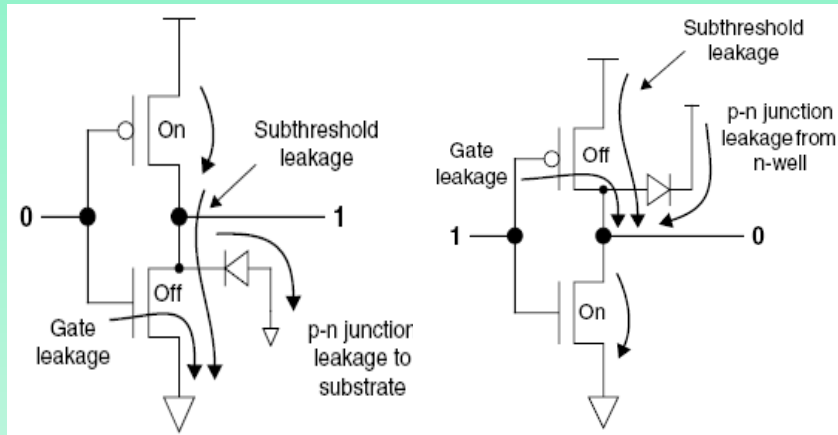
## Capacitance include:

- 1) **Output node capacitance of the logic gate:** due to the drain diffusion region.
- 2) **Total interconnects capacitance (wire capacitance):** has higher effect as technology node shrinks.
- 3) **Input node capacitance of the driven gate:** due to the gate oxide capacitance.

# Power Consumption

$$\text{Static power } (P_{\text{static}} = N_t I_L V)$$

## 1). Diode reverse bias–I1



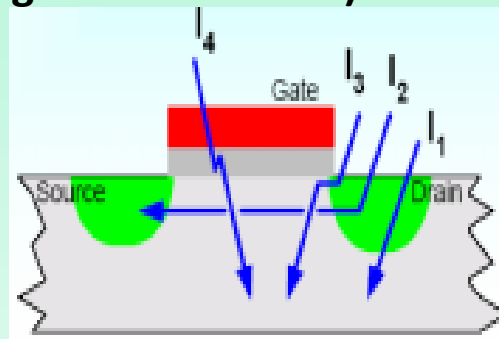
## 2). Sub threshold current – I2

$V_{gs} < \sim V_{th} \rightarrow$  carrier diffusion causes sub threshold leakage.

- ☐  $V_{gs} \leq 0 \rightarrow$  accumulation mode.
- ☐  $0 < V_{gs} \ll V_{th} \rightarrow$  depletion mode.
- ☐  $V_{gs} \sim V_{th} \rightarrow$  weak inversion.
- ☐  $V_{gs} > V_{th} \rightarrow$  Inversion.

## 3). Gate induced drain leakage – I3

- Higher supply voltage.
- thinner oxide.
- increase in  $V_{db}$  and  $V_{dg}$ .



## 4). Gate oxide tunneling – I4

- high electric field across a thin gate oxide.
- Direct tunneling through the silicon oxide layer if it is less than 3–4 nm thick.

Fig 2:leakage currents

## Question 1

Which factor we can adjust in order to reduce the power consumption of a circuit?

$$P = A N t C_L V^2 f + N t I_L V$$



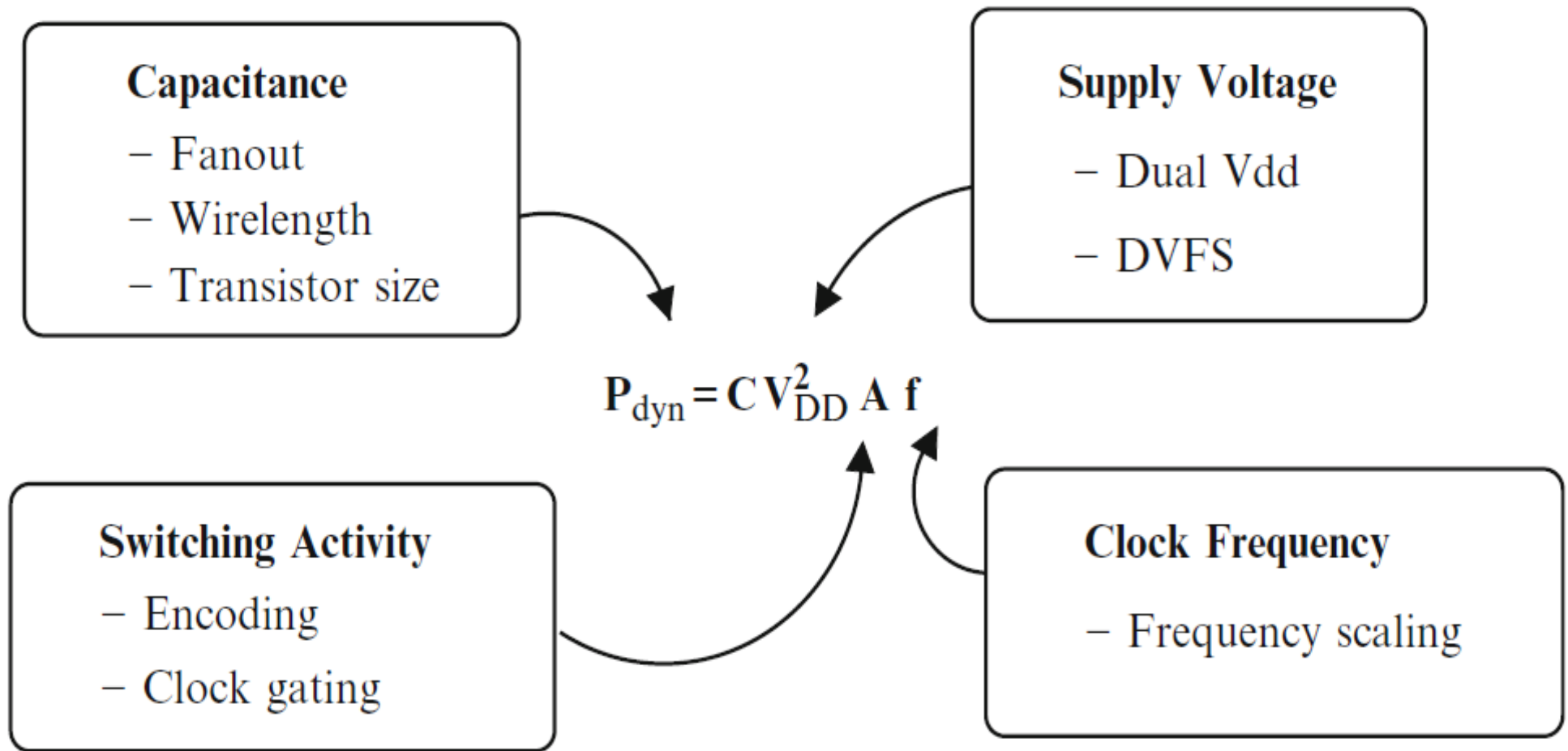
# Contents

- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power

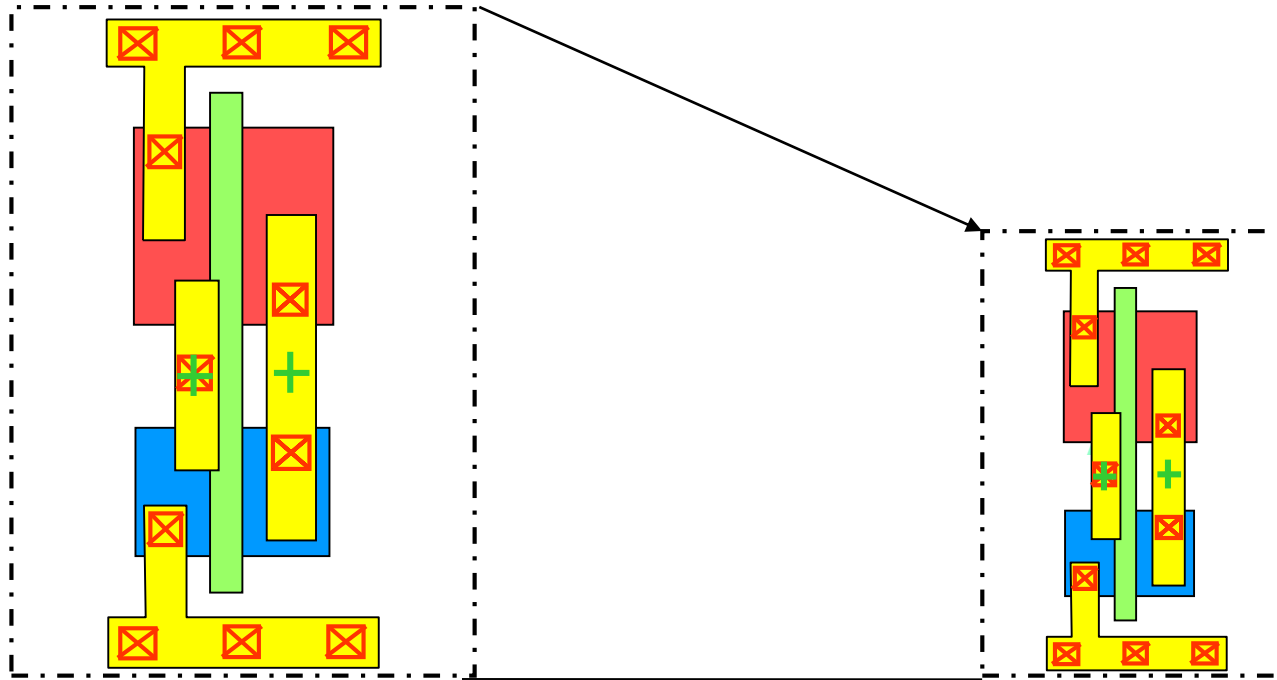
# Contents

- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power
  - ❖ Dynamic Power
  - ❖ Leakage Power

# Techniques for reducing Dynamic Power



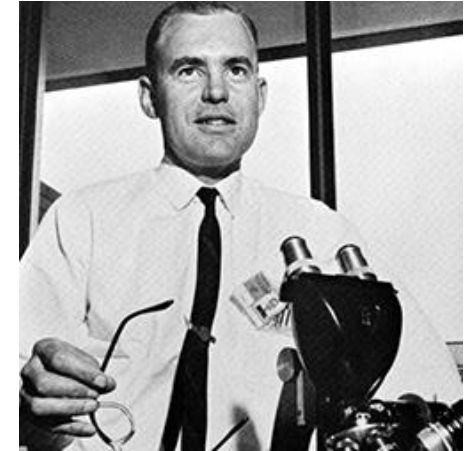
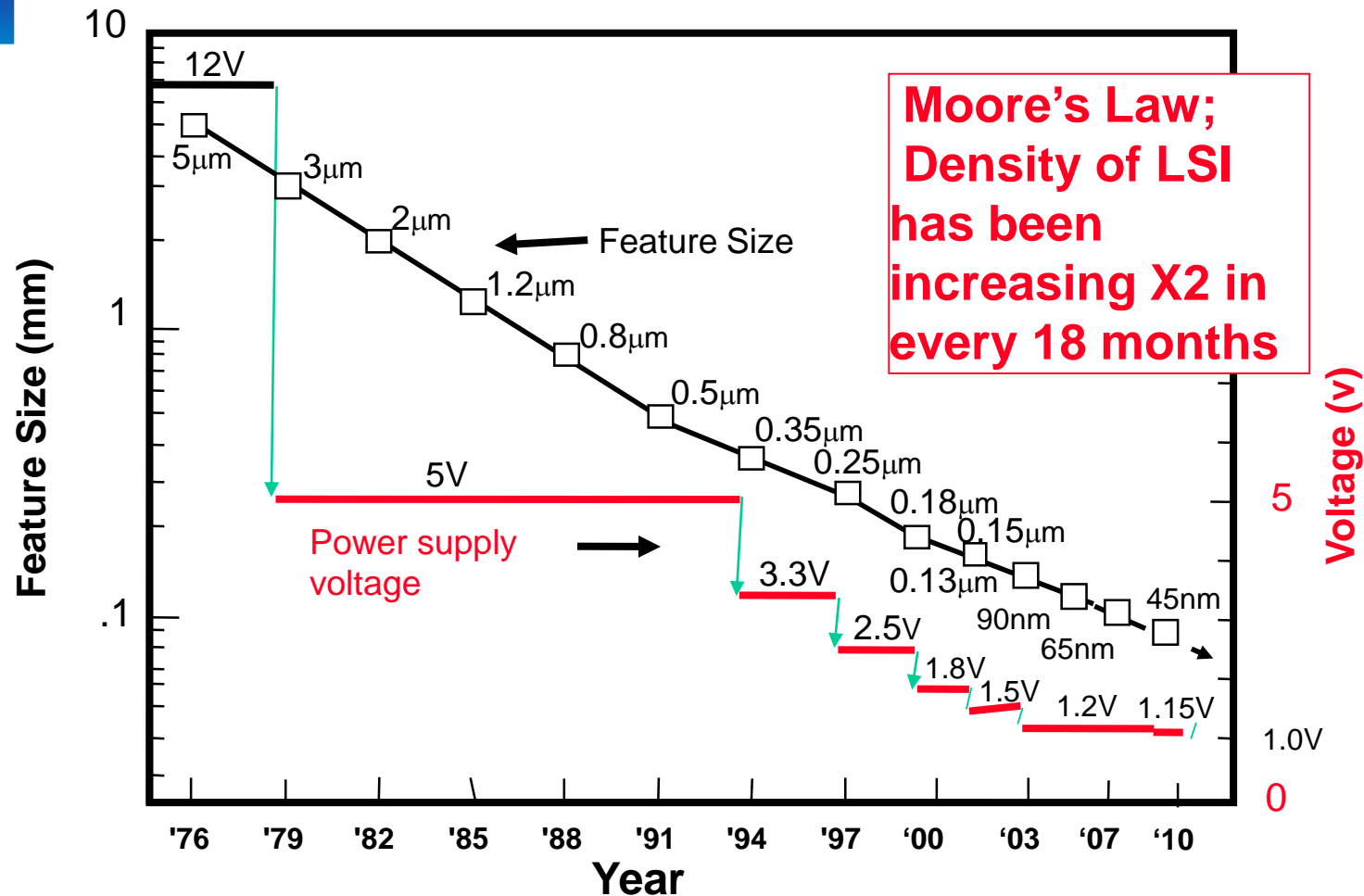
# Transistor Scaling



Small feature size (Almost equivalent to gate length)

→ **small gate capacitance and small wire capacitance**

# Moore's law



*Gordon Moore,  
co-founder of Intel Corp.*

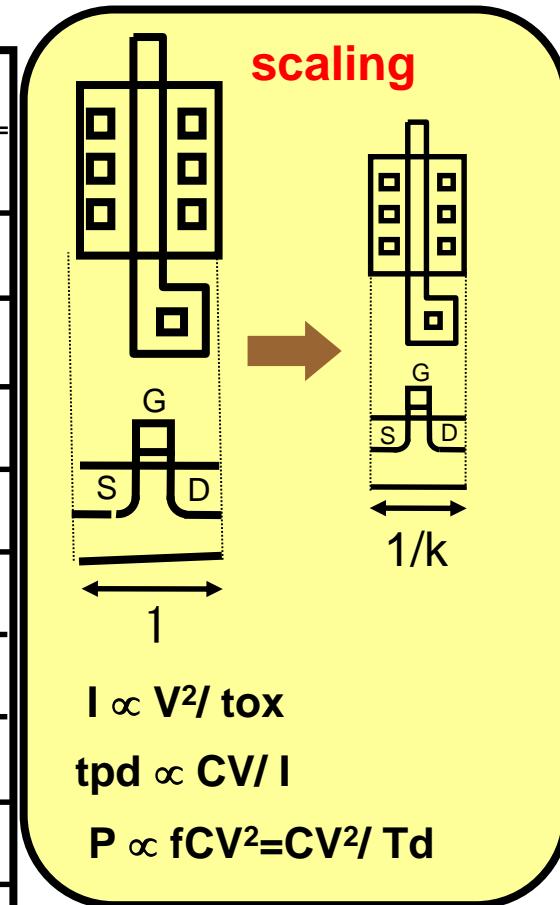
After 0.35 μm generation, power supply voltage has been reduced to maintain electric field constant in gate oxide.

After 90nm generation, power supply voltage has been saturated not to increase leakage current

# Scaling Law

**k: scaling factor**

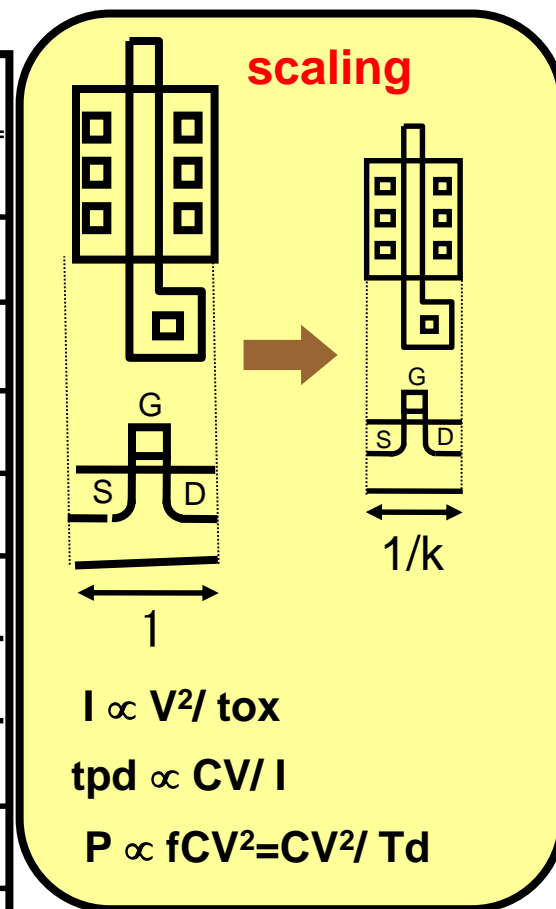
	voltage constant (very Old)	electric field Constant (old)	Voltage and Tox Constant (now)
One side of length (X)	$1/k$	$1/k$	$1/k$
Number of gate/unit area	$k^2$	$k^2$	$k^2$
Gate oxide thickness (tox)	$1/k$	$1/k$	1
Power supply voltage (V)	1	$1/k$	1
Electric field strength (E)	$k$	1	$k$
Saturation current (I)	$k$	$1/k$	1
Capacity (C)	$1/k$	$1/k$	$1/k^2$
Delay time (tpd)	$1/k^2$	$1/k$	$1/k^2$
Power consumption of one gate (P)	$k$	$1/k^2$	1
Power density (P/X <sup>2</sup> )	$k^3$	1	$k^2$



# Scaling Law

**k: scaling factor**

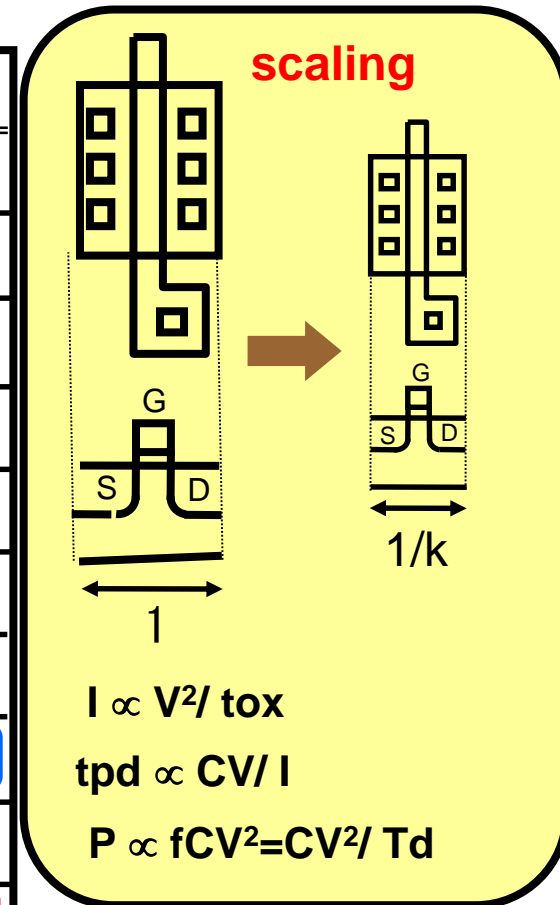
	voltage constant (very Old)	electric field Constant (old)	Voltage and Tox Constant (now)
One side of length (X)	$1/k$	$1/k$	$1/k$
Number of gate/unit area	$k^2$	$k^2$	$k^2$
Gate oxide thickness (tox)	$1/k$	$1/k$	1
Power supply voltage (V)	1	$1/k$	1
Electric field strength (E)	$k$	1	$k$
Saturation current (I)	$k$	$1/k$	1
Capacity (C)	$1/k$	$1/k$	$1/k^2$
Delay time (tpd)	$1/k^2$	$1/k$	$1/k^2$
Power consumption of one gate (P)	$k$	$1/k^2$	1
Power density (P/X <sup>2</sup> )	$k^3$	1	$k^2$



# Scaling Law

**k: scaling factor**

	voltage constant (very Old)	electric field Constant (old)	Voltage and Tox Constant (now)
One side of length (X)	$1/k$	$1/k$	$1/k$
Number of gate/unit area	$k^2$	$k^2$	$k^2$
Gate oxide thickness (tox)	$1/k$	$1/k$	1
Power supply voltage (V)	1	$1/k$	1
Electric field strength (E)	$k$	1	1
Saturation current (I)	$k$	$1/k$	1
Capacity (C)	$1/k$	$1/k$	$1/k^2$
Delay time (tpd)	$1/k^2$	$1/k$	$1/k^2$
Power consumption of one gate (P)	$k$	$1/k^2$	1
Power density (P/X <sup>2</sup> )	$k^3$	1	$k^2$



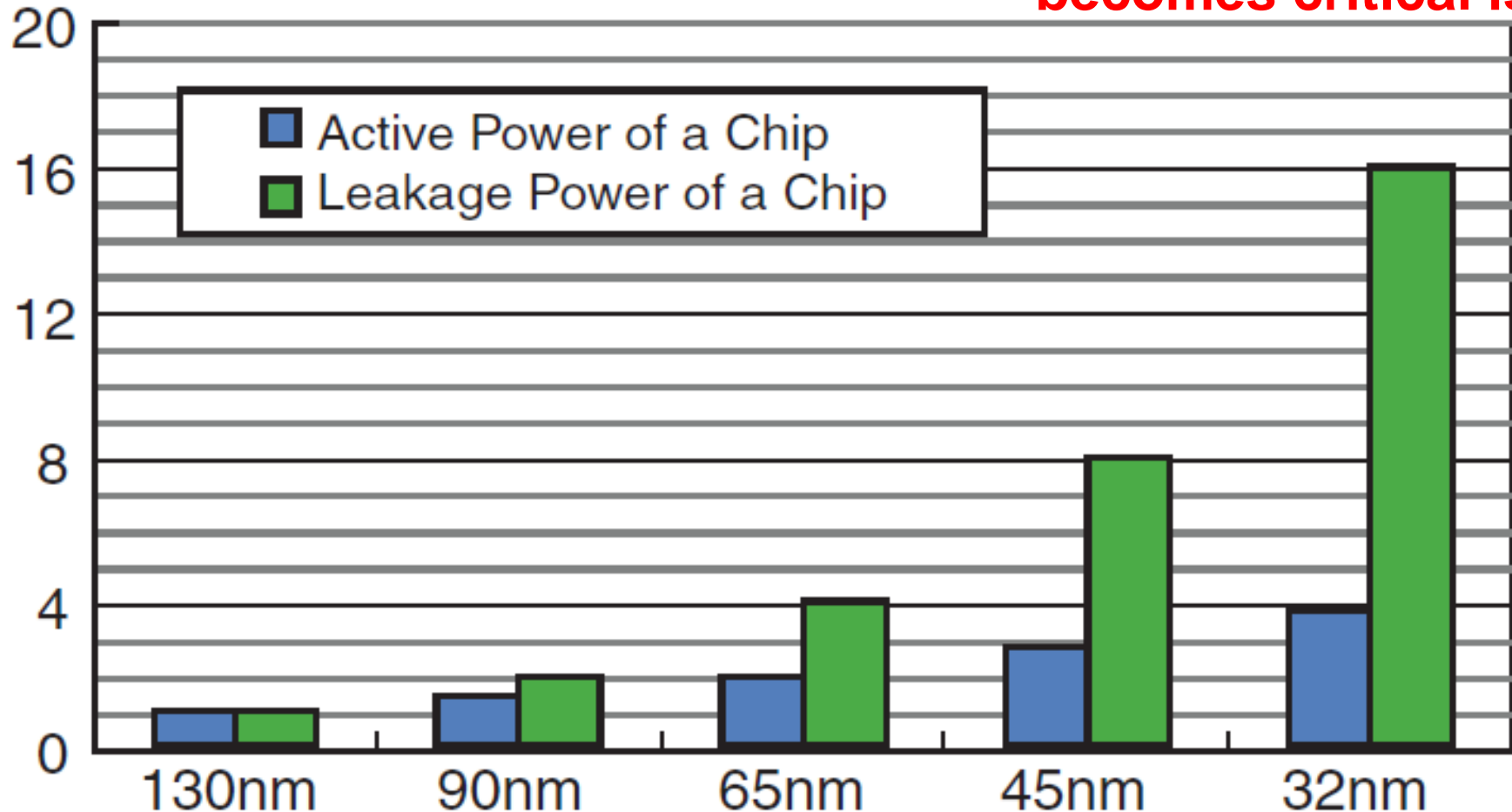
Success history of  
semiconductor  
development !!

Power density  
becomes serious  
problem



# Issue of Feature Scaling

**Leakage Power  
becomes critical issue**

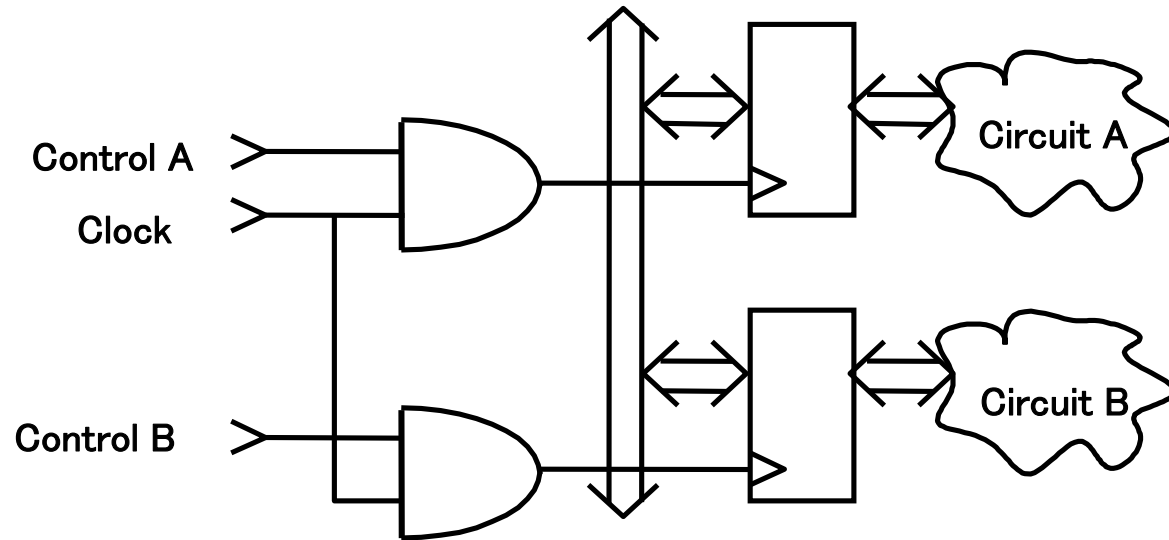


## Question 2

**Why don't we continue to reduce Supply voltage?**

# Clock Gating

$$P = \textcolor{red}{A} N t C_L V^2 f + N t I_L V$$



**Feature:** Inhibit clock signal of unused block.

**Effect:** AC power reduction (Small A)

**Status:** Representative and inevitable low power technique.

**Issue: ineffective to leakage**

# Clock Gating

## RX Family: Efforts to Reduce Power Consumption

The RX600 Series utilizes a 90 nm ultrafine process and a variety of technologies to reduce

### ▶ Reducing operating current consumption

#### ● 90 nm process

The 90 nm ultrafine process reduces the load capacitance (gates and wiring).

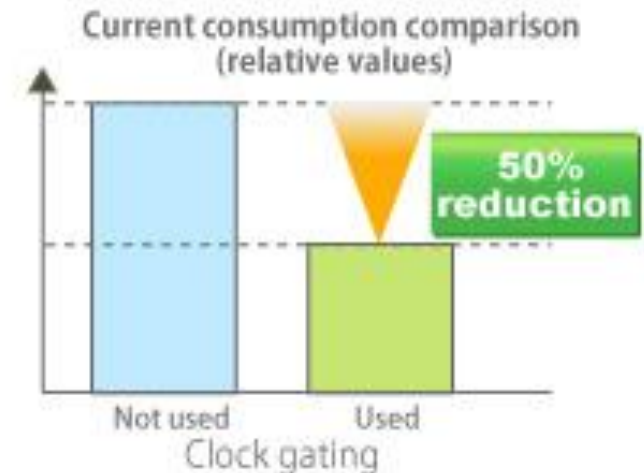
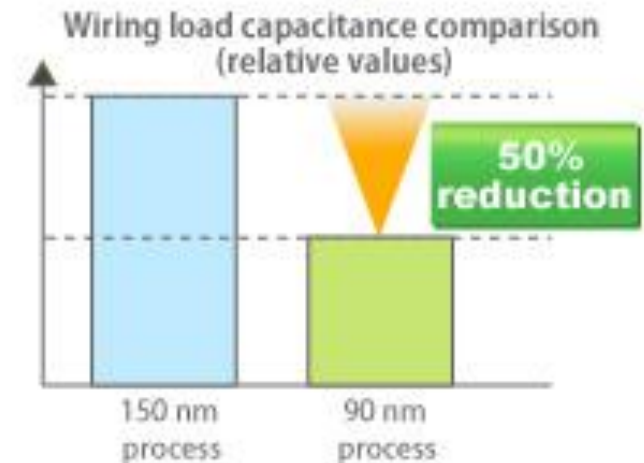
#### ● Clock gating technology

Analyzes the operation sequence and dynamically shuts off the clock supply to logic blocks that do not need it.

### ▶ Reducing standby current consumption

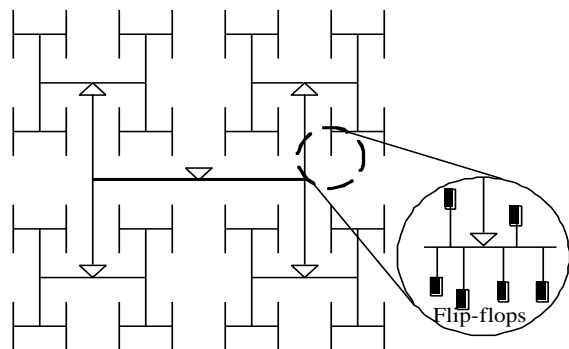
● Optimized use of both low-leak (high-V<sub>th</sub>) cells and high-speed (low-V<sub>th</sub>) cells for reduced standby current.

● Fine subdivision of power blocks in low-power mode to shut off power to inactive portions.



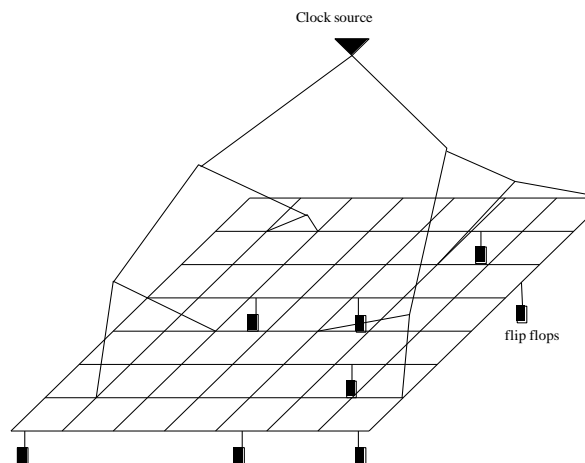
From <http://www.renesas.com>

# CTS - Classification



## Tree

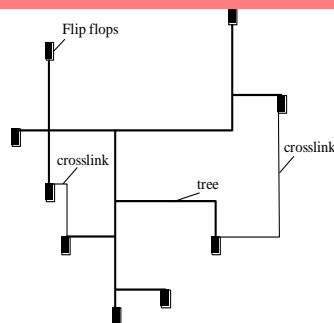
- low cost (wiring, power, cap)
- higher skew, jitter than mesh
- widely used in ASIC designs
- clock gating



## Mesh

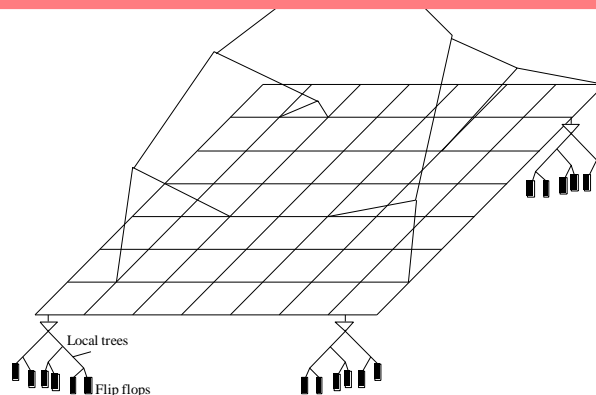
- excellent for low skew, jitter
- high power, area, capacitance
- difficult to analyze
- clock gating not easy
- used in modern processors

Best architecture depends on the application



## Hybrid: tree + cross-links

- low cost (wiring, power, cap)
- smaller skew, jitter than tree
- difficult to analyze



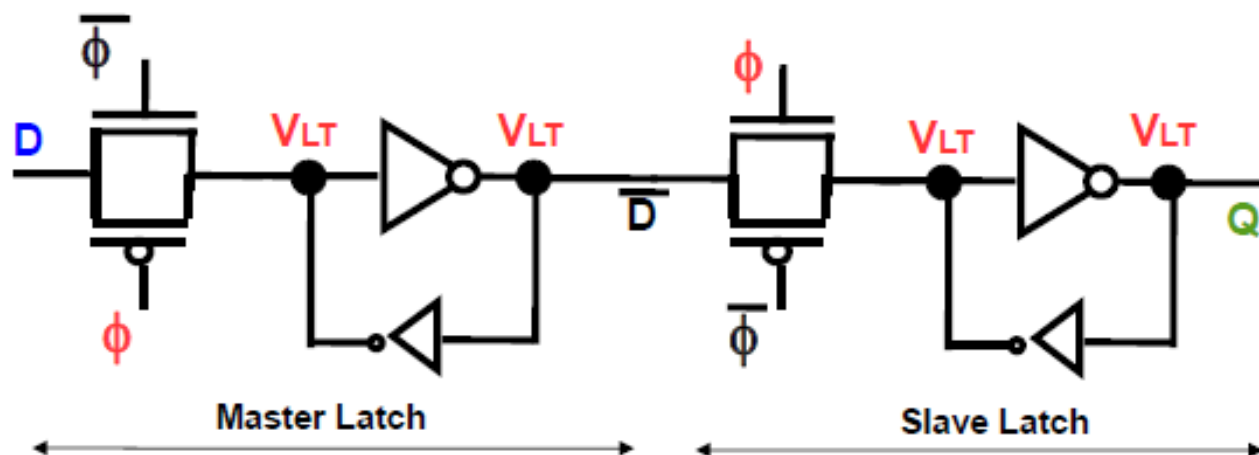
## Hybrid: mesh + local trees

- suitable for coarse mesh

# Clock Gating - Issues

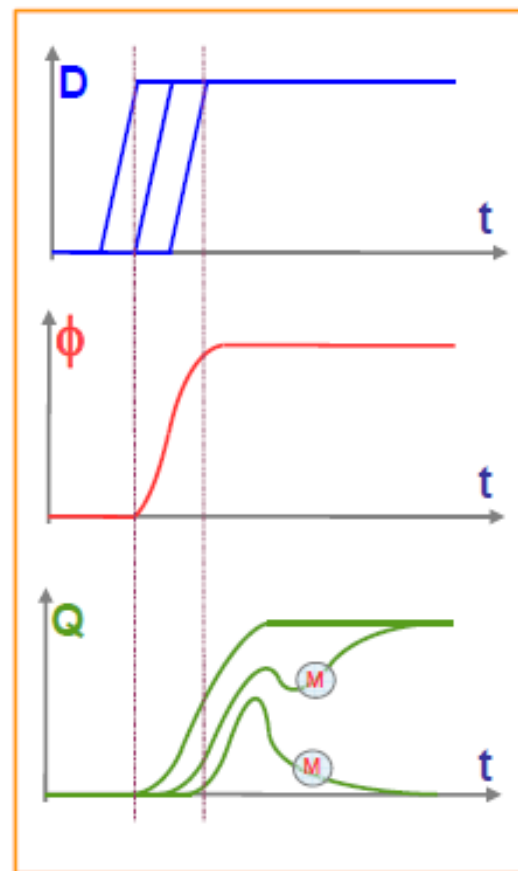
## Remind: Metastability

How does FF output (FF/Q) response while FF is in metastability?



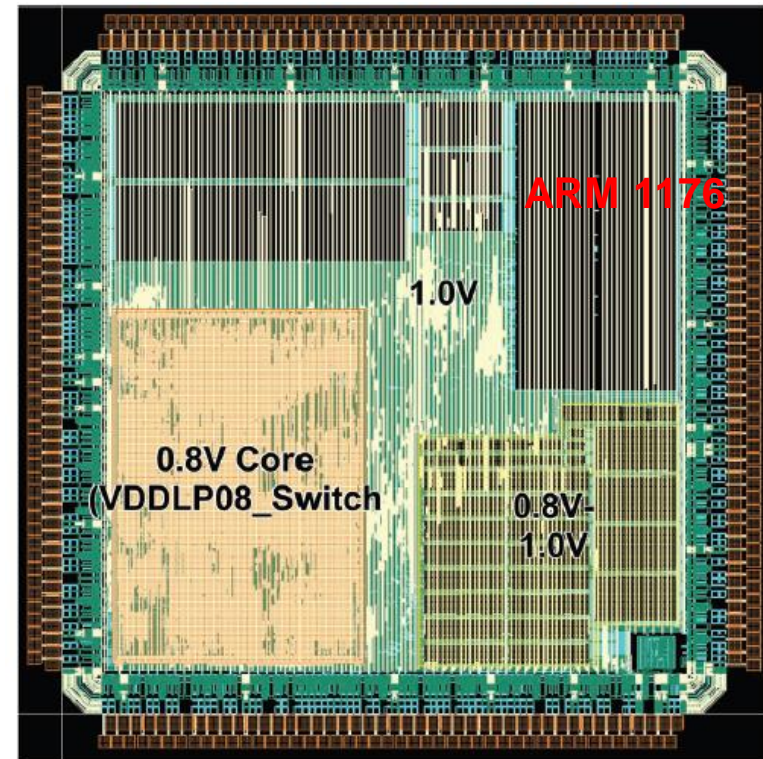
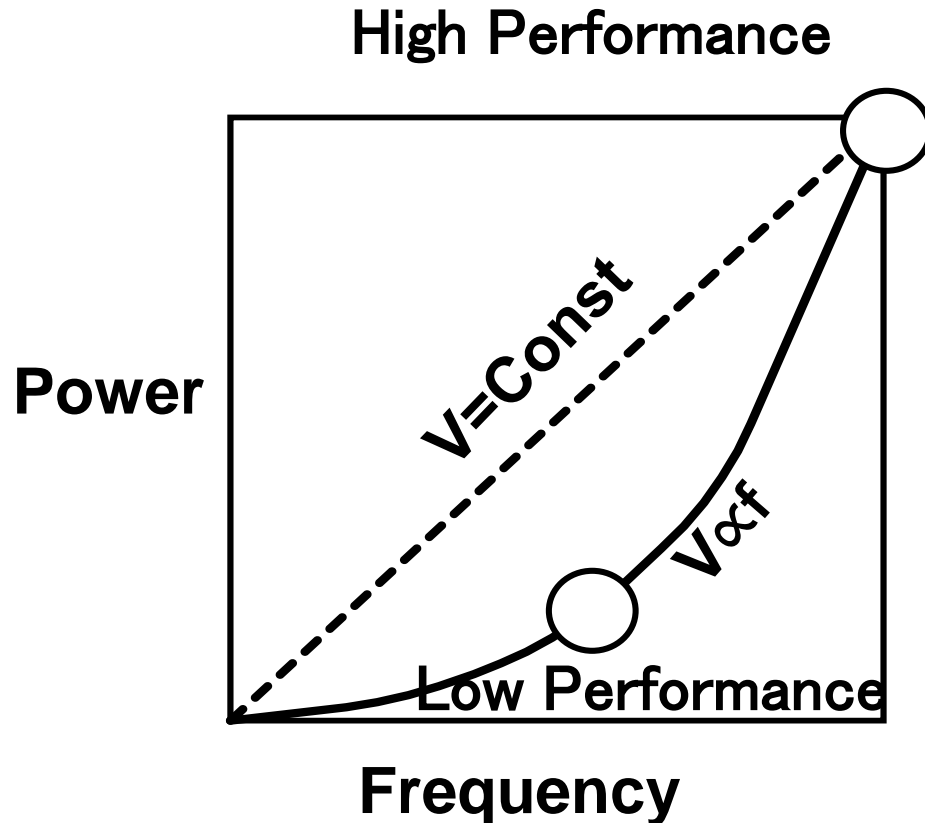
When  $\Phi$  is rising (and assume that master latch is entering to metastability), slave latch is active. Then, value of  $\overline{D}$  is transferred to the inverter loop. So, the FF output (FF/ $Q$ ) will be also  $V_{LT}$ .

FF/ $Q$  will keep this potential level for unpredictable time and the final value is also unpredictable as shown on waveform picture.



# DVFS

## – Dynamic Voltage Frequency Scaling –



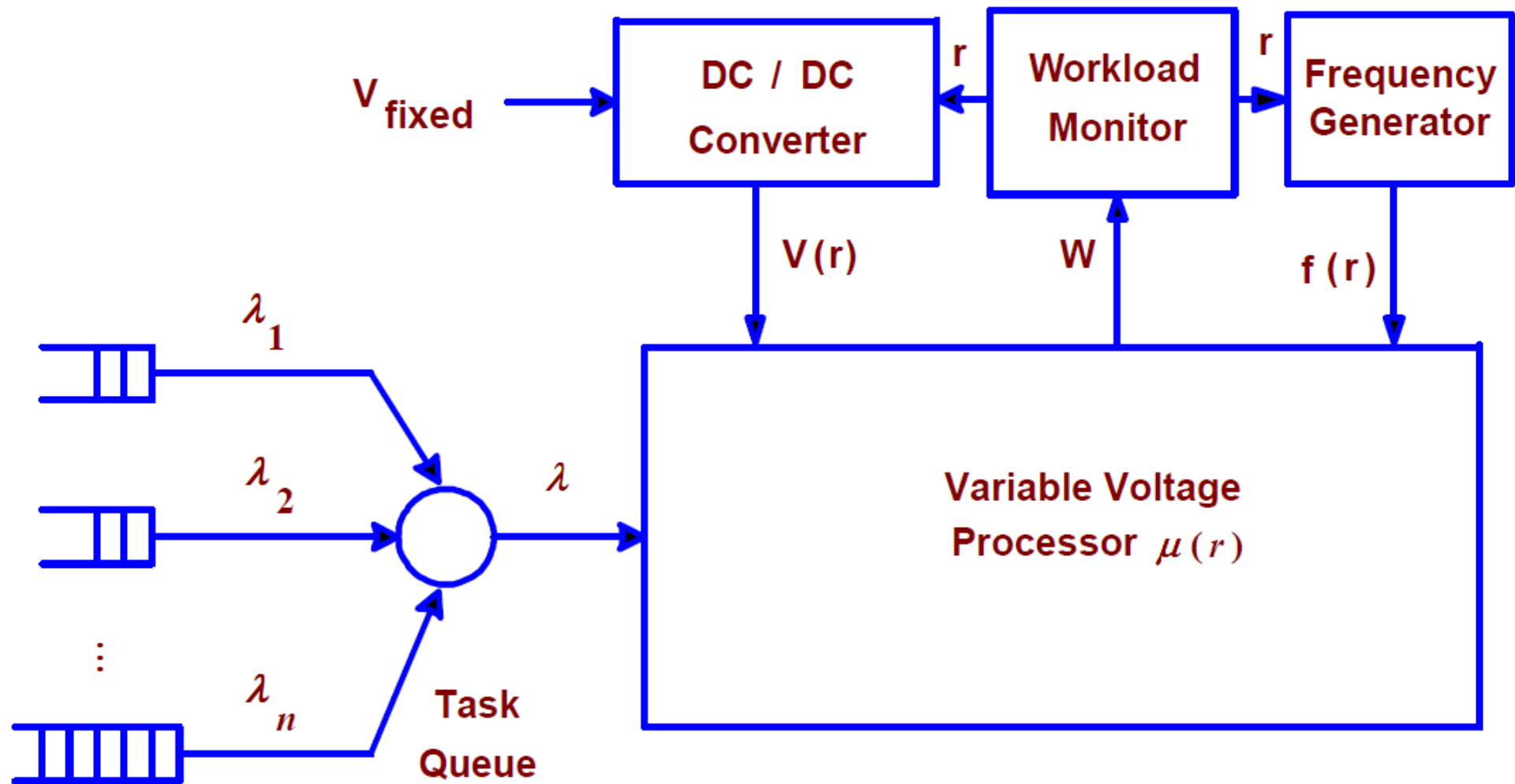
**Technique:** Scale the Supply voltage for each Block based on the operation frequency

**Effect:** can reduce power significantly; but impact to the complexity



# DVFS

## – Dynamic Voltage Frequency Scaling –



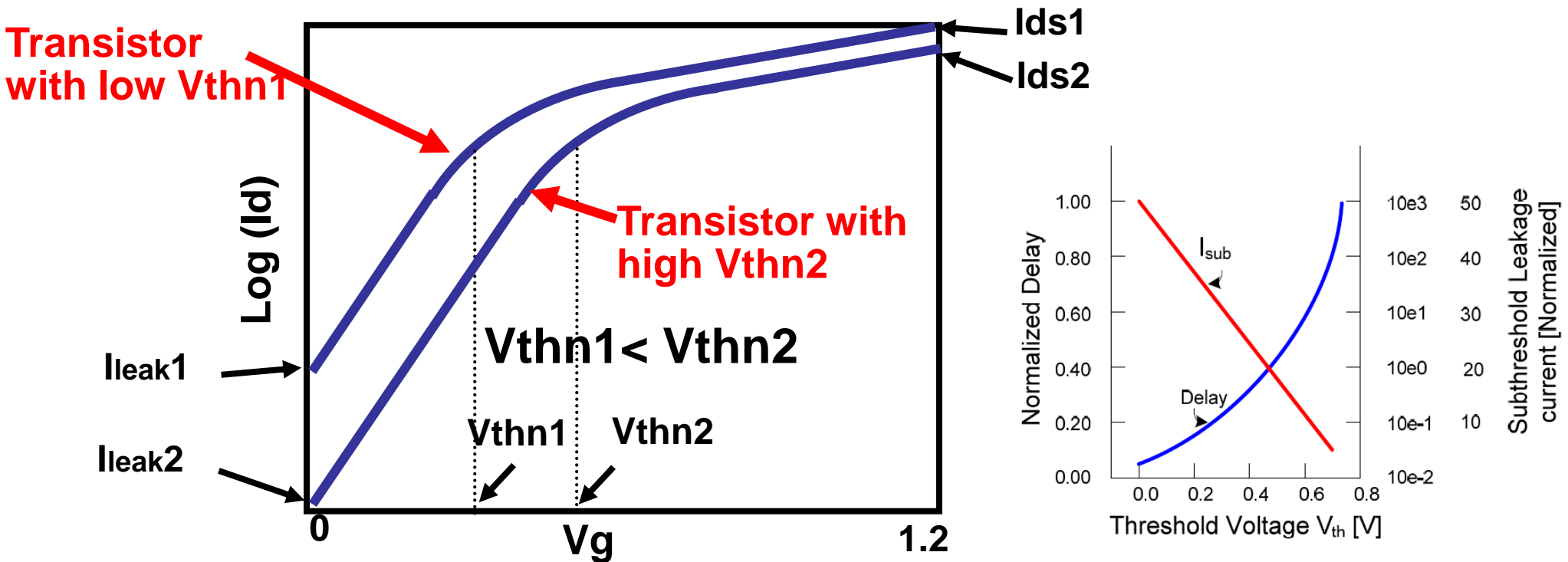


# Contents

- Low power SOC requirement
- CMOS device characteristics and power of LSI
- Scheme for Low Power
  - ❖ Dynamic Power
  - ❖ Leakage Power

# Multi VT Technique

Remind:  $V_{th}$  effect



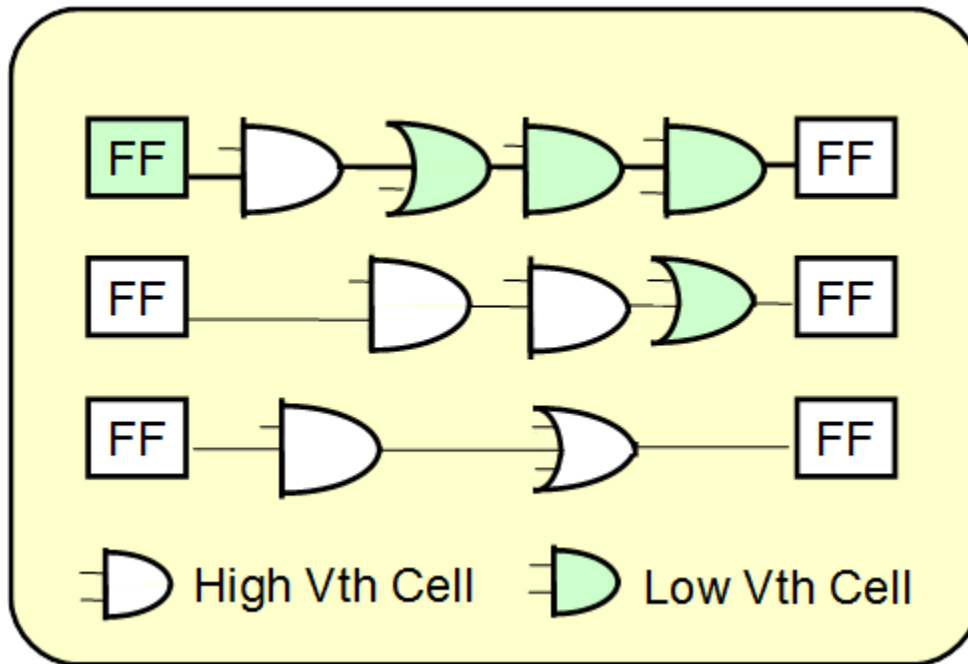
- ◆  $V_{th}$  can be tuned during fabrication by ion implantation.
- ◆ Transistor with lower  $V_{th}$  can flow large drain current at  $V_g = 1.2V$ , but has large leakage at  $V_g = 0V$  ( $I_{leak1} > I_{leak2}$ )

# Multi VT Technique

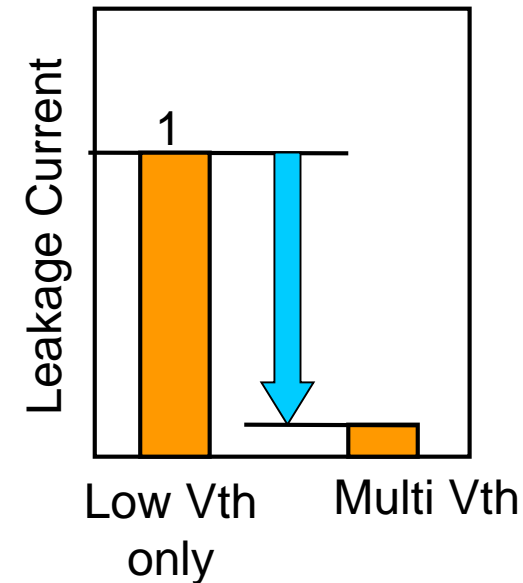
$$P = A N_t C_L V^2 f + N_{tH} I_{LH} V + N_{tL} I_{LL} V$$

$$I_{LH} < I_{LL}$$
$$N_t = N_{tH} + N_{tL}$$

*Structure*



*Effect*

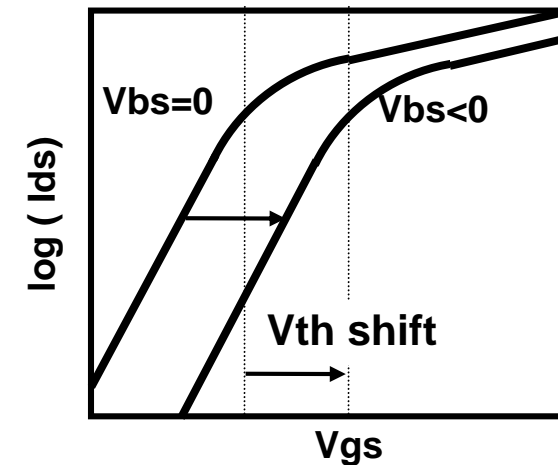
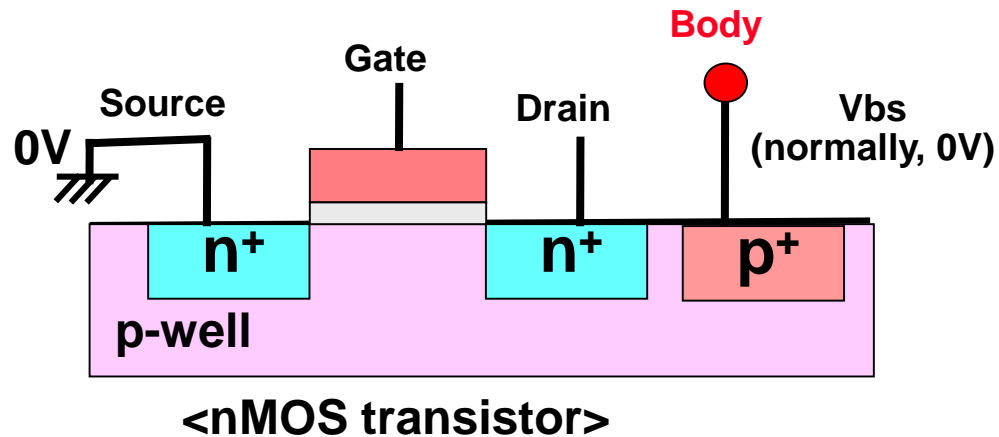


**Technique:** Use high and low Vth transistors and optimize the assignment of the Vth of cell.

**Effect:** Leakage can be reduced on both standby and operation mode.

# Static Body Bias Technique

## Remind: Body effect



< threshold voltage dependency on body voltage >

$$V_{thn} = V_{t0} + k(\sqrt{-V_{bs} + 2\phi} - \sqrt{2\phi})$$

$V_{t0}$ ,  $k$ ,  $\phi$ : constants

### Body Effect

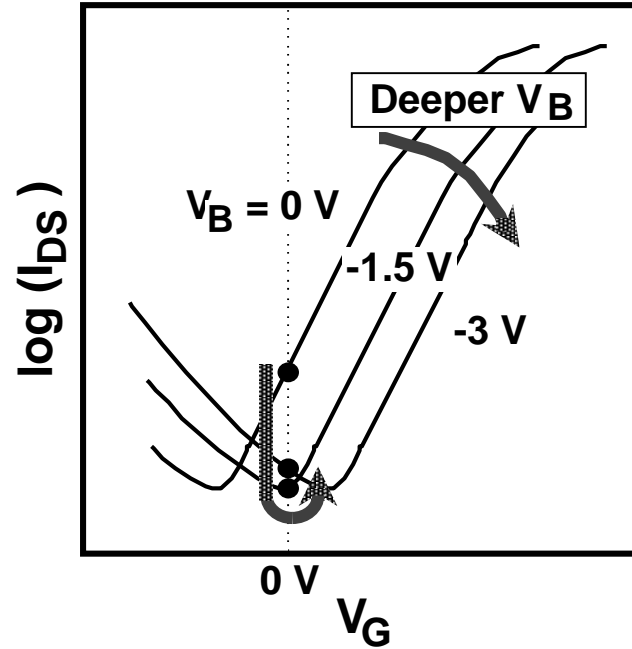
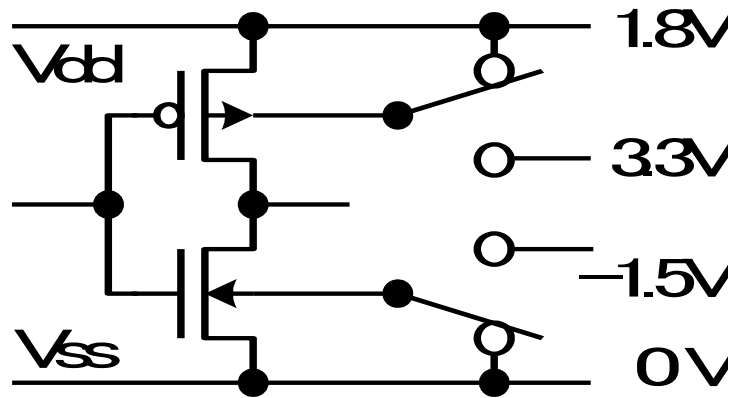
<nMOS>

- $V_{bs} < 0$ , then  $V_{thn}$  increases.
- low speed, low leakage

<pMOS>

- $V_{bs} > 0$ , then  $|V_{thp}|$  increases.
- low speed, low leakage

# Static Body Bias Technique



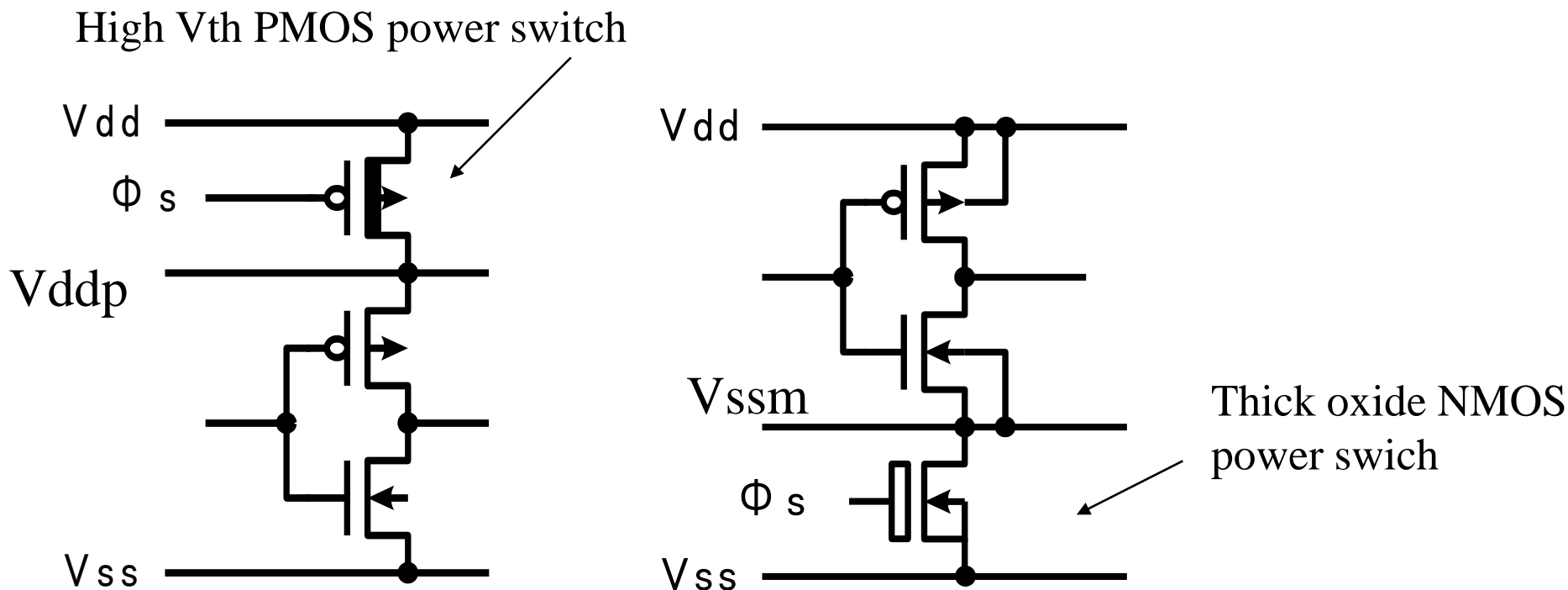
Negative body bias is applied at standby mode.

Effect: 1.5 – 3 orders magnitude leakage reduction

## Question 3

**What is the difference between “Multi Vth” technique and “Static Body bias” technique?**

# Power Gating - Concept



**Sleep transistor (Power switch) shut off the all leakage of the circuit**

**Effect: To decrease the leakage by 2 to 3 orders of magnitude**

**Issue: Data retention of FF**

# Power Gating – Classification

## ❑ Hierarchical:

- Fine-grain
- Coarse-grain

## ❑ Layout structure:

- Column structure
- Ring structure

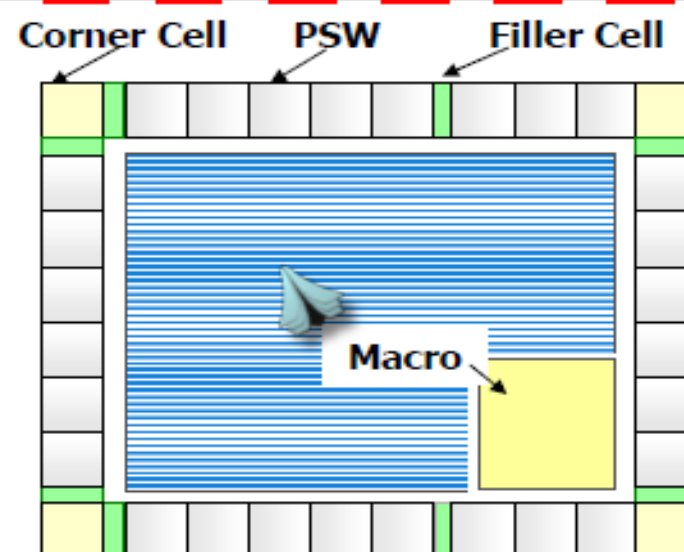
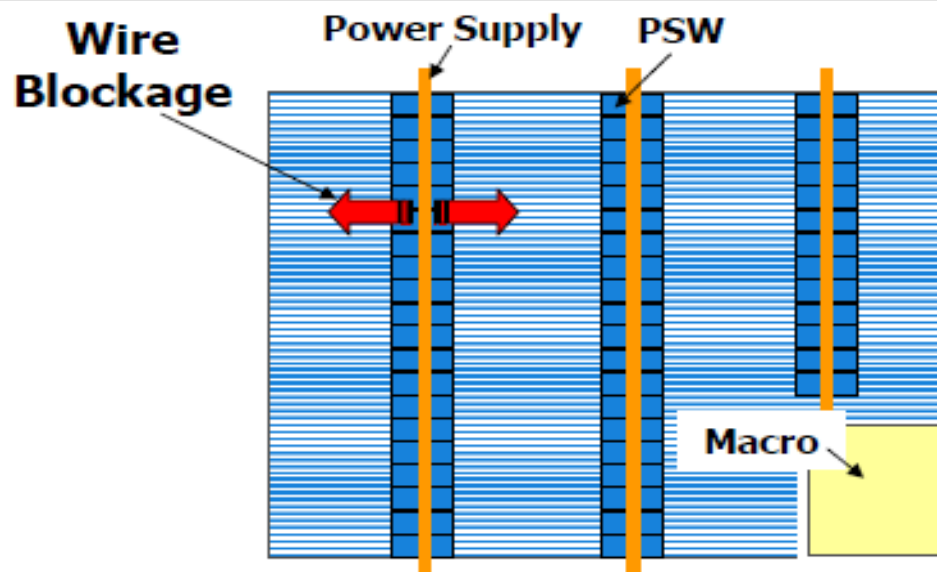


# Power Gating – Classification

	Fine-grain	Coarse-grain
Concept	<ul style="list-style-type: none"><li>▪ SW cells are applied to each Logic Gate</li></ul>	<ul style="list-style-type: none"><li>▪ SW cells are applied to each module</li></ul>
Power consumption reduction	<ul style="list-style-type: none"><li>▪ Good effect</li></ul>	<ul style="list-style-type: none"><li>▪ Less effect</li></ul>
Area overhead	<ul style="list-style-type: none"><li>▪ Large area</li></ul>	<ul style="list-style-type: none"><li>▪ Small area</li></ul>
Implementation	<ul style="list-style-type: none"><li>▪ Complicated</li><li>▪ Difficult to re-use</li></ul>	<ul style="list-style-type: none"><li>▪ Easy</li><li>▪ Easy to re-use</li></ul>

# Power Gating – Classification

	COLUMN Type PSW	RING Type PSW
Area efficiency	■ Tr size is small	◆ Tr size is large
Routing efficiency	■ PSW prevents horizontal routing	◆ No routing restriction inside Power Domain
Hard IP	■ Not applicable ■ Need to re-create physical design for power shut-off	◆ Easy to re-use for power shut-off



# Summary (1/2) – basic concept

- Low power has become important issue in recent LSI.
- Many techniques have been proposed and developed to reduce the power of LSI.
- Each technique has its own benefits and issues. Designers **MUST combine many techniques to maximize the benefits and minimize the side effects.**
- Formula to calculate power consumption

$$P = A Nt C_L V^2 f + Nt I_L V$$

# Summary (2/2) – Low power techniques

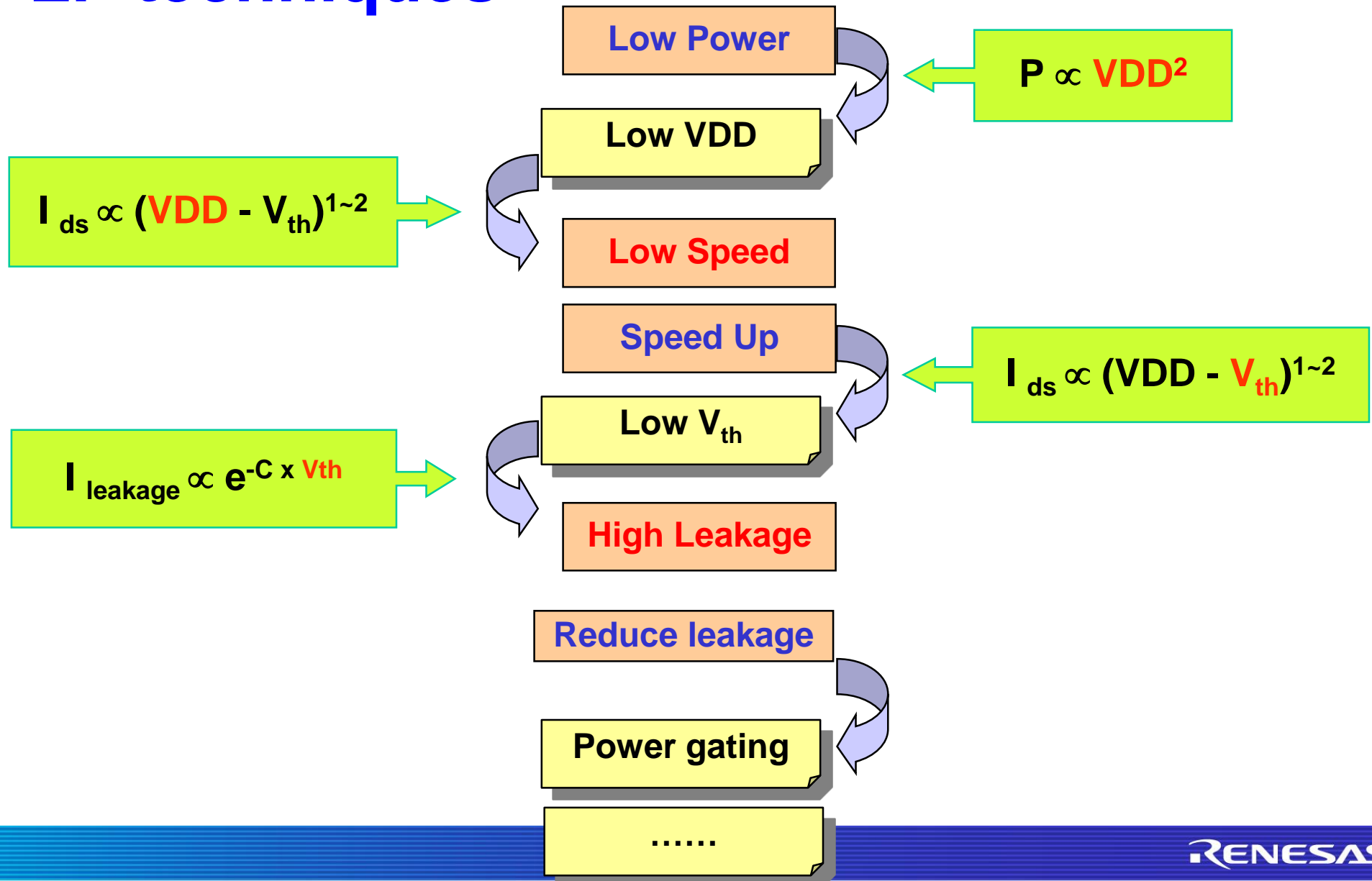
## Trade-offs for low-power techniques

Techniques, such as dynamic frequency scaling, require a sophisticated methodology

	Leakage power	Dynamic power	Timing	Area penalty	Methodology impact	Methodology change
Low-power optimization	10%	10%	0%	10%	None	None
Multi- $V_t$	6X	0%	0%	0%	Low	Multi- $V_t$ library needed
Clock gating	0%	20%	0%	<2%	Low	Clock-gating cells needed and extra overhead in STA
Multisupply voltage	2X	40% to 50%	0%	<10%	Medium	Microarchitecture and methodology needs to be domain-aware; Need voltage regulators and level shifters; verification and analysis challenge
Power shutoff	10X to 50X	0%	4% to 8%	5% to 15%	Medium-high	Insertion of switch cells; retention flops; wake-up and shutdown time analysis
Dynamic voltage frequency scaling	2X to 3X	40% to 70%	0%	<10%	High	Multimode optimization and analysis flow needed: Clock synchronization
Substrate biasing	10X	—	10%	<10%	High	Maintain well separation; multiple power rail distribution; timing analysis

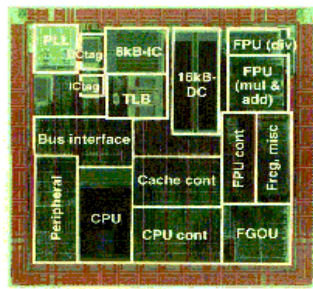
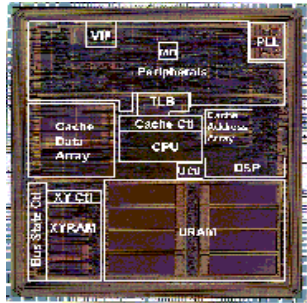
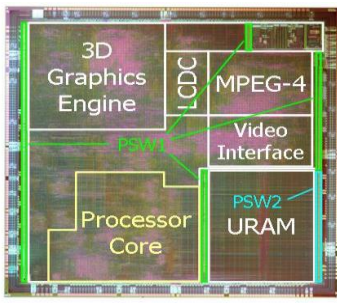
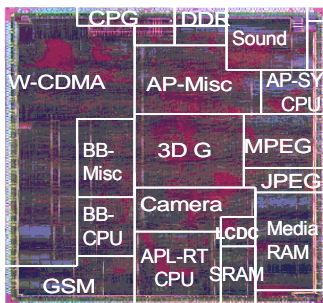
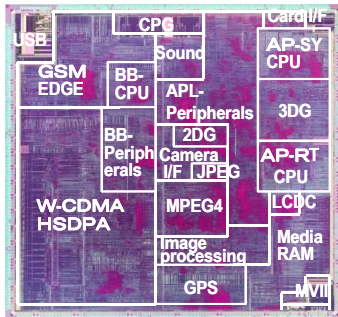
Source: Cadence Design Systems

# Example – Relationship between LP techniques



# Example - Renesas LP products



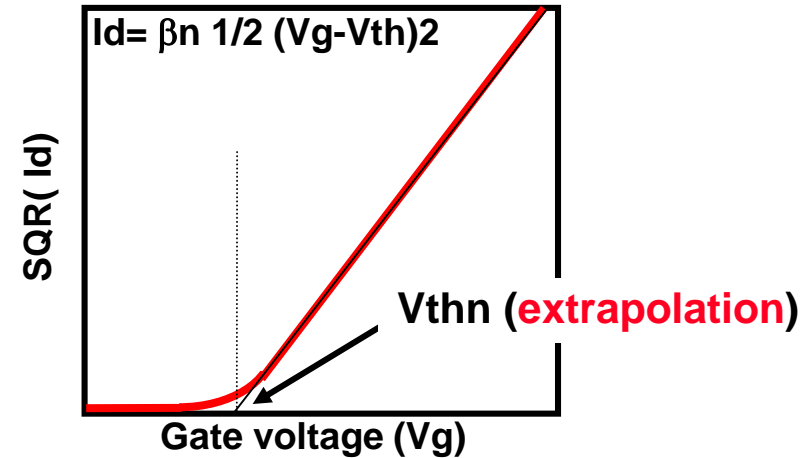
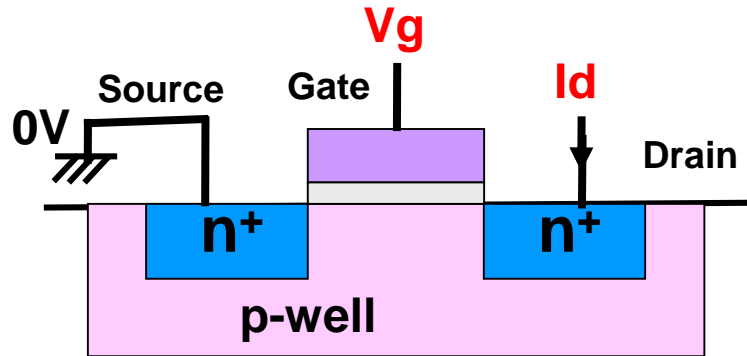
ISSCC98	ISSCC02	ISSCC04	ISSCC06	ISSCC07
				
0.25um	0.18um	0.13um	90nm	90nm
SH-4	SH-Mobile1	SH-Mobile3	SH-MobileG1	SH-MobileG2
Clock Gear Standby Back-bias	Dual Vth $\mu$ I/O On chip SRAM U-standby	Pointer pipeline Activation Control Resume Standby	Hierarchical Power Domain Control	Triple-Vth Core-Standby Dynamic Module Stop for Bus



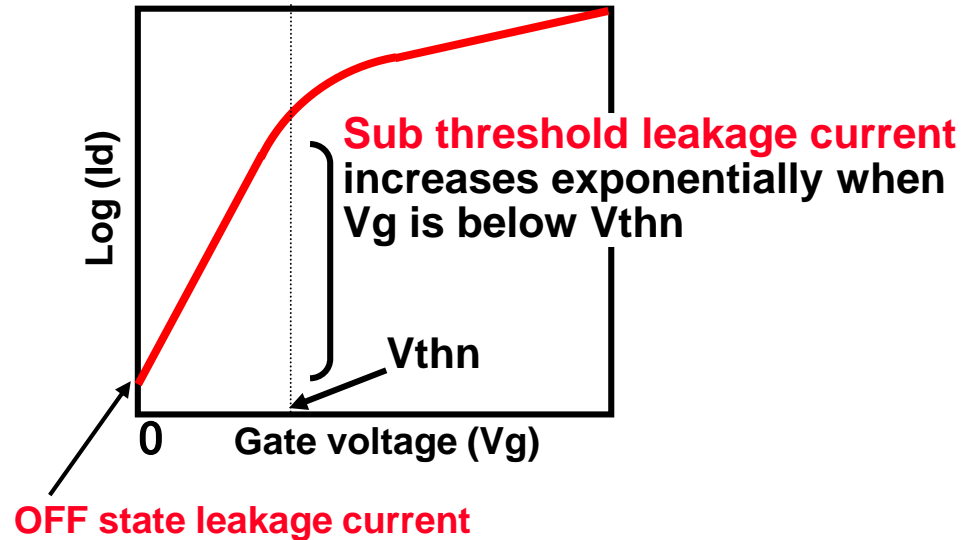
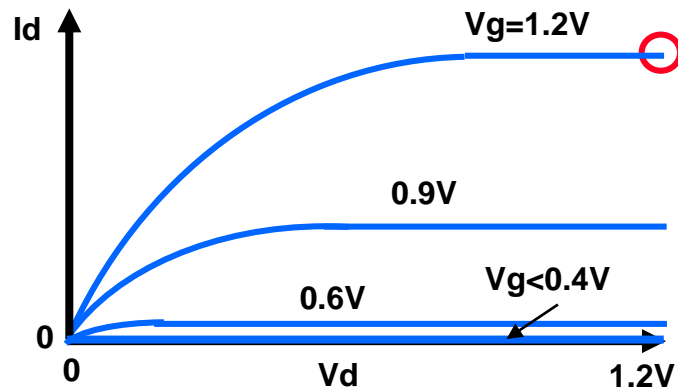
**RENESAS**

Renesas Design Vietnam Co., Ltd.

# Threshold Voltage of NMOS ( $V_{thn}$ )



<Characteristics of drain current>





# Clock Gating - Issues

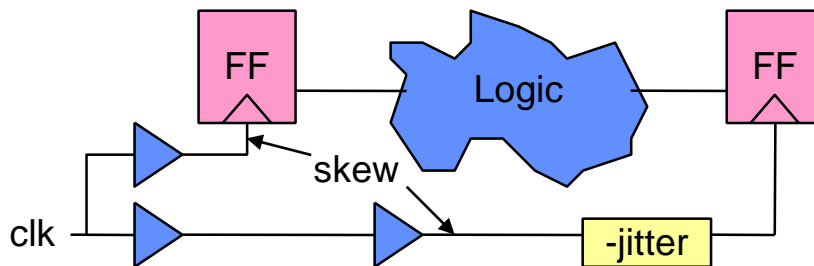
## Clock *skew*

The deterministic (**knowable**) difference in clock arrival times at each flip-flop  
Caused mainly by **imperfect balancing of clock tree/mesh**

## Clock *jitter*

The random (**unknowable**) difference in clock arrival times at each flip-flop  
Caused by on-die process, Vdd, temperature variation, PLL jitter, crosstalk,  
Static timing analysis (STA) accuracy, layout parameter extraction (LPE)  
accuracy

Clock uncertainty  $\Delta \equiv \text{skew} \pm \text{jitter}$



# Formula

Dynamic Power  $P_{switching} = C_{switching} \cdot VDD^2 \cdot f$

Average Short Circuit Current

$$I_{SC} = \frac{\beta \cdot \tau_{in}}{12 \cdot VDD} \cdot (VDD - 2V_{th})^3 \cdot f$$

gain\_factor:  $\beta_n = \beta_p = \beta$ ,

Threshold\_Voltage:  $V_{thn} = |V_{thp}| = V_{th}$

Sub-threshold Leakage Current

$$I_{DS} = K \cdot e^{(V_{GS} - V_{th}) \cdot q / nkT} \cdot (1 - e^{-V_{DS} \cdot q / kT})$$

$K$ : function of technology,  $V_{GS}$ : gate – to – source voltage,  $V_{DS}$ : drain – to – source voltage,

$V_{th}$ : threshold voltage,  $q$ : electronic charge,  $k$ : Boltzmann constance,  $T$ : temperature,

$n$ : nonlinearity constance 1 ~ 2 , ( $kT \cong 0.0259$ )

## Scaling Relationship

Quality	Cons field Scaling	Constant Voltage Scaling
Gate Capacitance	$C'_g = C_g / S$	$C'_g = C_g / S$
Drain Current	$I'_D = I_D / S$	$I'_D = I_D . S$
Power Dissipation	$P' = P / S^2$	$P' = P . S$
Power Density	$P' / Area' = (P / Area)$	$P' / Area' = S^3 P / Area$
Delay	$t'_d = t_d / S$	$t'_d = t_d / S^2$
Energy	$E' = \frac{P}{S^2} \times \frac{t_d}{S} = \frac{P.t_d}{S^3} = \frac{1}{S^3} E$	$E' = E / S$

# Threshold Voltage of NMOS ( $V_{thn}$ )

