



# **Green SoC**



### VLSI SYSTEM LAB, YONSEI University School of Electrical & Electronic Engineering



## Contents



- 1. Introduction
- 2. Power classification
- 3. Power performance relationship
- 4. Low power design
  - 1. Architecture and algorithm level
  - 2. Block and logic level
  - 3. Circuit level
  - 4. Device level
- 5. Summary



# Introduction





# **Technology Scaling**



### Technology scaling : Moore's law

 The number of transistors that can be placed on an integrated circuit has doubled approximately every 18 months



# **Development Trend**





### Scaling (More Moore)

- More devices are integrated in a chip
- New scaling road map
  - Not only 'geometrical scaling' for 2D device, but also 'equivalent scaling' for 3D device
- Beyond bulk CMOS
   FinFET, SOI…
- Functional diversification (More than Moore)
  - Several functions are merged in a chip





# SoC Performance





### SoC performance : exponentially increase!!

• Thanks to both device technology and design methodology





### **SoC Power Consumption Problem**



### **♦** SoC power consumption : 'also' severely increase

• After 15 years, x10 power is required...







# SoC Power Density Problem



### Power density : exponentially increase!!

- Power consumption per die area (W/cm<sup>2</sup>)
- We would soon reach power densities of nuclear power plants or rocket nozzles in a few years!!





# **Process Variation Problem**

### Process variation : Result of scaling

- Global variation and local variation
  - Global variation
    - Comes from fabrication, lot, wafer processes
    - Different process corner (NMOS-PMOS : SS/SF/TT/FS/FF)
  - Local variation
    - Truly random variation between device with identical layout





## **Process Variation Problem**



### Performance variation due to process variation

- Frequency difference  $\approx 30\%$
- Leakage current difference  $\approx$  x20
- $\Rightarrow$  Process variation should be considered in SoC design







- $\bullet ~ I_{\rm D} \propto ~ W/L^* (V_{\rm DD} \text{-} V_{\rm TH})^{\alpha}$
- $V_{TH}$  variation  $\Rightarrow$   $I_D$  variation  $\Rightarrow$  Performance Variation !!
- Need more design margin due to process variation  $\Rightarrow$  V\_{DD}  $\uparrow$

### Yield limitation

• Because of process variation, failure probability  $\uparrow \Rightarrow$  Yield  $\downarrow$ 







#### 2009 ITRS SPECIAL TOPICS

### ENERGY

Energy consumption has become an increasingly important topic of public discussion in recent years because of global CO<sub>2</sub> emission. Since semiconductor electronics are broadly applicable to energy collection, conversion, storage, transmission, and consumption/usage, it is not surprising that the ITRS addresses many factors of significance to energy issues. In general, the ITRS documents the impressive trends and, more importantly, sets aggressive targets for future electronics energy efficiency, for example, computational energy/operation (per logic and per memory-bit state changes). The most detailed targets relate directly to semiconductor materials, process, and device technologies, which form the bases of integrated-circuit manufacturing and components, respectively.

### ♦ Green SoC

- = Low power SoC design !!!
- = Low process variation (high yield) design



# **Power Classification**



# **Power Classification**



Power consumption of CMOS circuits

$$P_{total} = P_{dynamic} + P_{static}$$

 $P_{dynamic} = P_{sw} + P_{sc}$ 



# Switching Power





 $I=C_L dV/dt=C_L \Delta V f$ 

 $\mathsf{P}_{\mathsf{sw}} {=} \mathsf{IV}_{\mathsf{DD}} {=} \mathsf{C}_{\mathsf{L}} \Delta \mathsf{V} \; \mathsf{V}_{\mathsf{DD}} \mathsf{f}$ 

In digital circuit,  $\Delta V=V_{DD}$ 

 $\mathsf{P}_{\mathsf{sw}} {=} \mathsf{IV}_{\mathsf{DD}} {=} \mathsf{C}_{\mathsf{L}} \mathsf{V}_{\mathsf{DD}}{}^2 \mathsf{f}$ 

P<sub>sw</sub> is due to the charge and discharge (output transition) of the capacitors driven by the circuit according to input transition.

• 
$$P_{sw} = C_L V_{DD}^2 f$$



# **Short Circuit Power**





P<sub>sc</sub> is caused by the simultaneous conductance of PMOS and NMOS during input and output transitions.

• 
$$P_{sc} = (\beta/12)(V_{DD}-2V_{TH})^3 (t_3-t_1)$$





# Static Power : P<sub>sub</sub>, P<sub>gate</sub> & P<sub>junc</sub>



# Power – Performance Relationship







# Power consumption equation P<sub>sw</sub> = C<sub>L</sub>V<sub>DD</sub><sup>2</sup>f

• 
$$P_{sc} = (\beta/12) (V_{DD} - 2V_{TH})^3 (t_3 - t_1)$$

- $P_{sub} \propto Exp[(V_{GS}-V_{TH})/mv_{T}] V_{DD}$
- $P_{gate} \propto WL (V_{GS}/T_{OX})^2 V_{DD}$
- $P_{junc} \propto Exp[V_D/v_T 1] V_{DD}$

### • Case.1 : $V_{DD} \downarrow$

- All power consumption  $\downarrow$
- However...
  - $\bigstar \text{Delay} \propto \ C_L V_{\text{DD}} / I_{\text{D}} \propto \ C_L V_{\text{DD}} / (V_{\text{DD}} \text{-} V_{\text{TH}})^{\alpha}$
  - $\clubsuit \text{ If } V_{\text{DD}} \downarrow, \text{ Delay} \uparrow$
  - $\Rightarrow$  Performance loss





# **V**<sub>DD</sub> Scaling Limitation

### $\blacklozenge$ Low V\_{DD} limitation with process variation

- $V_{DD.min} = V_{T0} + K\sigma(V_T)$ •  $\sigma(V_T)$ : 1-sigma of  $V_T$  variation  $\succ \propto T_{ox}N_A^{0.25}(LW)^{-0.5}$
- Significant increment of  $\sigma(V_T)$  with technology scaling (LW $\downarrow\downarrow\downarrow$ )
- $\Rightarrow$  V<sub>DD</sub> scaling meets the limitation!!
- ⇒ Process variation tolerant circuit design technique is required!!



[7] K.Itoh, "Adaptive Circuits for the 0.5-V Nanoscale CMOS Era", ISSCC, 2009



High V<sub>TH</sub>



Power consumption equation
 P<sub>sw</sub> = C<sub>L</sub>V<sub>DD</sub><sup>2</sup>f

• 
$$P_{sc} = (\beta/12) (V_{DD} - 2V_{TH})^3 (t_3 - t_1)$$

- $P_{sub} \propto Exp[(V_{GS}-V_{TH})/mv_{T}] V_{DD}$
- $\mathbf{P}_{gate} \propto \mathbf{WL} \ (\mathbf{V}_{GS}/\mathbf{T}_{OX})^2 \ \mathbf{V}_{DD}$
- $P_{junc} \propto Exp[V_D/v_T 1] V_{DD}$

### ♦ Case.2 : V<sub>TH</sub> ↑

- $\bullet~\mathsf{P}_{\mathsf{sc}}\downarrow$  and especially,  $\mathsf{P}_{\mathsf{sub}}\downarrow$
- However...
  - $\bigstar \text{Delay} \propto \ C_L V_{\text{DD}} / I_{\text{D}} \propto \ C_L V_{\text{DD}} / (V_{\text{DD}} \text{-} V_{\text{TH}})^{\alpha}$
  - $\clubsuit \text{ If } V_{\text{TH}} \uparrow, \text{ Delay} \uparrow$
  - $\Rightarrow$ Performance loss



# Low Frequency





### **♦** Case.3 : f ↓

- $\mathsf{P}_{\mathsf{sw}}\downarrow$
- However...
  - $\clubsuit Throughput \propto f$
  - $\Rightarrow$ Performance loss



## Tradeoff





- $\Rightarrow$  Tradeoff between low power and high performance
- $\Rightarrow$  Low power design :
  - power reduction without performance degradation



# Low power design





## Low Power Design Methodology

### ◆ To make low power SoC...

- Architecture and algorithm levels
   Parallelism, Pipeline ...
- Block and logic levels
  - V<sub>DD</sub> / Frequency scheduling by monitoring workload (AVFS)
  - Temperature management to reduce leakage current
- Circuit level
  - ✤ Circuit type (Dynamic, static, …)
  - ✤ Circuit technique (Dual V<sub>DD</sub>, Dual V<sub>TH</sub>, MTCMOS, …
- Device level
  - Control the process parameter
    - ➤ Halo doping, retrograde well...
  - ✤ Low leakage new device
    - ➢ SOI, FinFET …



### Architecture and Algorithm Levels



## Parallelism





*N*: # of parallelism  $\delta$ : a slight increase in capacitance due to the extra routing

[8] A.P. Chandrakasan, "Minimizing power consumption in digital CMOS circuits", Proc. of IEEE,995



# **Pipeline**









< Pipeline implementation>

$$P_{pipe} = C_{pipe} V_{pipe}^2 f_{pipe}$$

N: # of pipeline stage  $\delta$ : a slight increase in capacitance due to the extra latch

28 [8] A.P. Chandrakasan, "Minimizing power consumption in digital CMOS circuits", Proc. of IEEE,995



# **Circuit level**





### **Circuit Level Low Power Techniques**

### Low power techniques

- Multiple channel length
- Stacked transistor
- Dual V<sub>DD</sub>
- Dual V<sub>TH</sub>
- MTCMOS (Multi Threshold voltage CMOS)
- DVS (Dynamic Voltage Scaling) : open-loop / closed loop



## **Critical Path**





### Critical Path : The worst case delay path

- Determines SoC's maximum performance
- # of critical path << # of non-critical path
- Fast non-critical path is just wasteful...
  - ⇒By increasing non-critical path's delay, we may achieve power reduction because of tradeoff relation between power & performance





# **Multiple Channel Length**

### Threshold voltage roll-off

- Longer L
  - ✤ Higher Vt
  - Low leakage with low performance
  - Used in non-critical path





# **Stacked Transistor**



♦ V<sub>M</sub> level
• V<sub>M</sub> > 0 due to leakage current.
• Negative V<sub>GS\_MN1</sub>
• Positive V<sub>SB\_MN1</sub>
→ Increase in V<sub>TH</sub> by body effect
P<sub>sub</sub> ≈  $\left(e^{\frac{-(V_{gs}-V_{th})}{mv_T}}\right)V_{dd}$ → Large reduction in I<sub>sub</sub>

Primary input vector control to utilize the stack effect in the standby mode

|        | High Vt | Low Vt |   |
|--------|---------|--------|---|
| 2 NMOS | 10.7X   | 9.96X  | Ì |
| 3 NMOS | 21.1X   | 18.8X  | ļ |
| 4 NMOS | 31.5X   | 26.7X  |   |
| 2 PMOS | 8.6X    | 7.9X   |   |
| 3 PMOS | 16.1X   | 13.7X  | Ì |
| 4 PMOS | 23.1X   | 18.7X  |   |



# Dual V<sub>DD</sub>



### Basic idea

• V<sub>DDL</sub>

✤ Logic gates off the critical path

- V<sub>DDH</sub>
  - ✤ Logic gate on the critical path
- Reduce power without degrading the performance









## **Dual V<sub>DD</sub> : Design Issue & Target**

### ♦ Issue

- Static current flow at a V<sub>DDH</sub> gate if it is directly drive by a V<sub>DDL</sub> gate
- Level converter is needed
- $\Rightarrow$  Overhead of area and power



### Design target

 For a give circuit, choose gates for V<sub>DDL</sub> application to minimize power consumption while maintaining performance with consider level converter.



# Dual V<sub>TH</sub> Voltages



#### ♦ HVt

- Assigned to transistors in noncritical path.
- Leakage saving in both standby and active modes
- ♦ LVt
  - Assigned to transistors in critical path
  - Maintained performance

# MTCMOS : Basic



#### MTCMOS : Multiple Threshold voltage CMOS Low power & low Energy • $E_{ToT} = E_{STD} + E_{ACT} = P_{static} * t_{STD} + P_{dynamic} * t_{ACT}$ • Portable device : $t_{STD} >> t_{ACT}$ 8.0 Basic circuit scheme 10E6 lsub tpd [nomalized] **Men** 10E5 • Two different Vt 6.0ីត្ 10E4 ✤ HVt (0.5~0.6V) 10E3 **❖** LVt (0.2~0.3V) 4.0 Vdd 10E2 • Two operating mode Delay Time 10E1 Active 2.010E0 978 Standby 10E-1 \_\_\_\_ 10E-2 0.8 0.0 0.6 0.20.00.4Threshold Voltage Vth [V]



# **MTCMOS : Scheme**



### Active mode

- SL=1 / <u>SL</u>=0
- $V_{DDV} \approx V_{DD} / V_{GNDV} \approx V_{GND}$
- LVt operating frequency

### Standby mode

- SL=0 / <u>SL</u>=1
- $V_{DDV} \& V_{GNDV}$  = floating
- HVt leakage







# MTCMOS : Constraint



### Performance constraint according to

- Normalized foot/head switch size : W<sub>H</sub>/W<sub>L</sub>
- Normalized cap on VDDV/VGNDV :  $C_V/C_O$

### Area penalty

 Relatively small because Head/Footswitches are shared by all logic gates on a chip (global foot switch)



# **DVFS : Basic Concept**

### Basic concept

- $P_{dynamic} = CV_{DD}^2 f$
- V<sub>DD</sub> and frequency scaling simultaneously
- V<sub>DD</sub> scaling
  - $\bigstar$  A best way to get low  $\mathsf{P}_{\text{dynamic}}$  because  $\mathsf{P}_{\text{dynamic}} \propto V_{\text{DD}}{}^2$
- Frequency scaling
  - Operating frequency = throughput
  - Not all task requires maximum throughput
  - By controlling the frequency, SoC improves energy efficiency





# DVFS : Open loop VS. Closed Loop



- Can not adapt to PVT variations
- Need more design margin
- Example
  - Enhanced SpeedStep technology of Intel



### Closed loop system

- Can adapt to PVT variations
- Need less design margin
- Example
  - Intelligent Energy Management technology of ARM
  - SmartReflex2 of TI OMAP processor





# DVFS (SONY, PDA)



### Block Diagram



Closed loop system

[12] M.Nakai, "Dynamic Voltage and Frequency Management for a Low-Power Embedded Microprocessor", JSSC, 2005 School of EEE



# **Delay Synthesizer Structure**



- Composed not only a simple transistor delay factor, but also wire delay and rise/fall delay
  - Gate delay component : one of nominal gate length and another of long gate length
  - RC delay component : wires from each of the four metal layers and its total length is 14mm





## **Delay Synthesizer Effect**





# **Operation (DVC+DFC)**



### Operation procedure



- Low → High : The main logic clock frequency is changed after the DVC confirms the voltage has increased enough
- High → Low : Both the DVC reference clock and the system clock are changed simultaneously



# **Device Level**



# Device Level Low Power Technique

### ♦ FinFET

- FinFET : Vertical structure
   Planar MOSFET width
   FinFET height
- σ(V<sub>T</sub>) ∝ T<sub>ox</sub>N<sub>A</sub><sup>0.25</sup>(LW)<sup>-0.5</sup>
   As scaling goes on, variation of planar MOSFET get worse
  - $\succ$  V<sub>DD</sub> scaling is impossible
  - - FinFET width doesn't occupy the active area
    - As scaling goes on, L\*W of FinFET can be maintained
    - >  $V_{DD}$  scaling is possible ⇒ low power !!



Planar MOSFET



FinFET





## Summary



Green SoC design  $\Rightarrow$  Low power & process variation tolerant SoC design  $\mathbf{P} = \mathbf{P}_{sw} + \mathbf{P}_{sc} + \mathbf{P}_{sub} + \mathbf{P}_{gate} + \mathbf{P}_{junc}$ P<sub>static</sub> P<sub>dynamic</sub> Power and performance : Trade-off Low power design • Architecture and algorithm level : parallelism, pipe line Block and logic level : workload monitoring, V<sub>DD</sub>/frequency scheduling Circuit level ♦ Long channel : Reduce  $I_{leak}$  by using  $V_{TH}$  roll off  $(V_{TH}\uparrow)$ Stacked MOSFET : Reduce I<sub>leak</sub> by using body effect (V<sub>TH</sub>↑) & negative V<sub>GS</sub> Dual V<sub>DD</sub>: Use low V<sub>DD</sub> at non-critical path • Dual  $V_{TH}$ : Use low  $V_{TH}$  at non-critical path ✤ MTCMOS: Use high V<sub>TH</sub> sleep TR (low leakage in stand-by mode) & low V<sub>TH</sub> logic (high performance in active mode) • DVFS : Reduce dynamic power by controlling both  $V_{DD}$  & frequency

• Device level : FinFET







# Study about advanced circuit level low power design technique

- SCCMOS (Super Cutoff CMOS) / SCCMOS-Zigzag
- SMTMOS (Selective MTCMOS: local footswitch)
- OZMTCMOS (Optimal Zigzag MTCMOS)
- BGCMOS (Boosted Gate CMOS)
- VTCMOS (Variable Threshold CMOS)
- DTCMOS (Dynamic Threshold CMOS)
- AVFS

