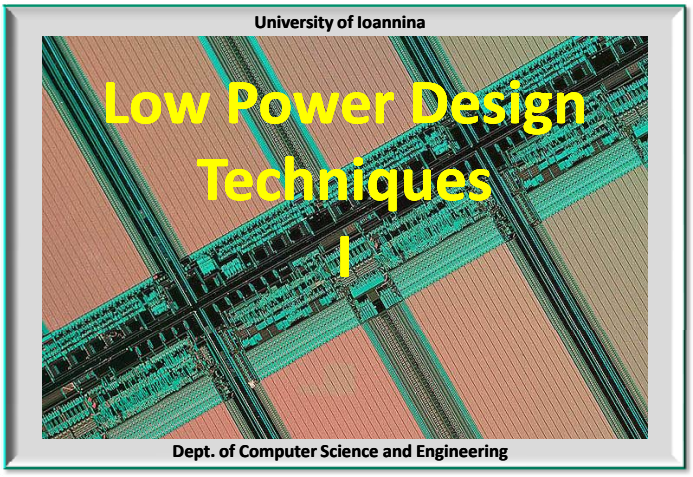
CMOS INTEGRATED CIRCUIT DESIGN TECHNIQUES



Y. Tsiatouhas

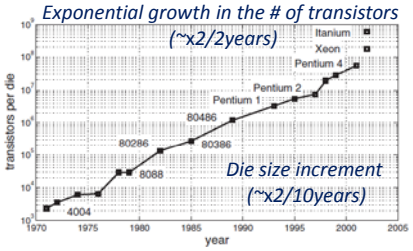
CMOS Integrated Circuit Design Techniques

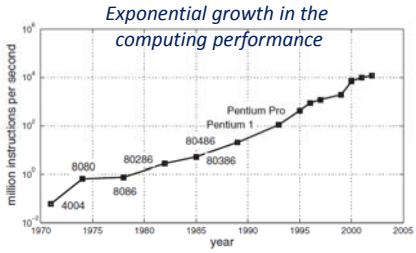
Overview

1. ***Power consumption sources***
2. ***Static power consumption reduction***

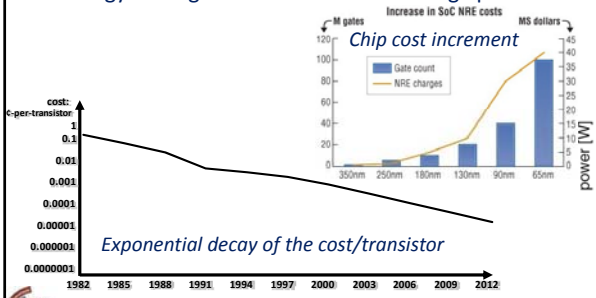
VLSI Systems
and Computer Architecture Lab

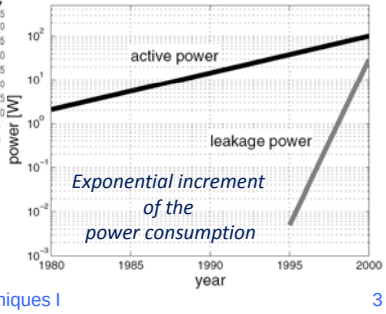
Transistor Scaling – Power – Performance





Technology Scaling ⇒ Increment of Switching Speed ⇒ Continuous Growth of Clock Frequencies





Power Consumption

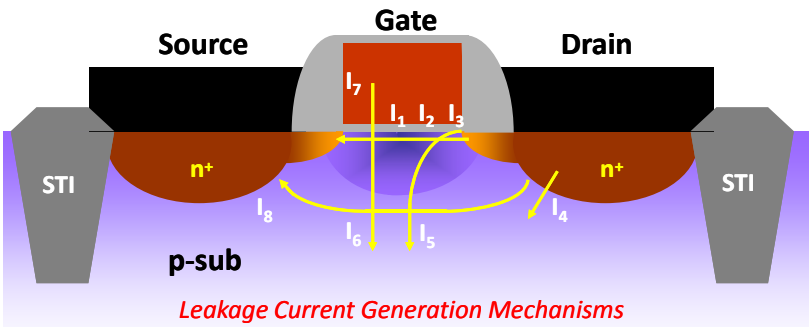
We distinguish three power consumption sources in digital integrated circuits:

- Static power consumption: due to leakage currents in the circuit, it remains even when the circuit is inactive (idle state).
- Dynamic power consumption: due to charging and discharging of circuit's internal node parasitic capacitances.
- Short-circuit power consumption: due to short-circuit currents between the power supplies during signal transitions at the inputs of the logic gates.

$$P = V_{DD} \cdot I_{leakage}(V_{DD}) + f_{clk} \cdot \alpha \cdot C_{tot} \cdot V_{DD}^2 + V_{DD} \cdot I_{sc}(V_{DD})$$

Static Power Consumption

Static Power Consumption



Leakage Current Generation Mechanisms

I_1: Weak Inversion
I_2: Drain-Induced Barrier Lowering
I_3: Narrow Width Effect
I_4: Reverse Bias Current
I_5: Gate-Induced Drain Leakage
I_6: Gate Oxide Tunnelling
I_7: Hot Carrier Injection
I_8: Punchthrough

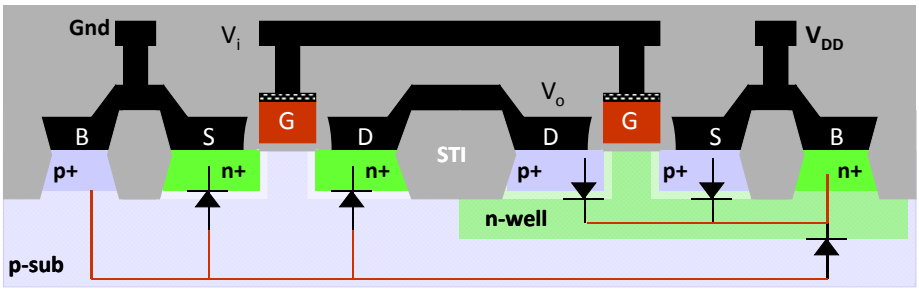
$$P_S = \sum_{1}^{n} Leakage\_Currents \times Power\_Supply = I_{off} \cdot V_{DD}$$

Static Power Consumption

Reverse Bias pn-Junction Leakage Current

1 *Static Power Consumption*

Inverter

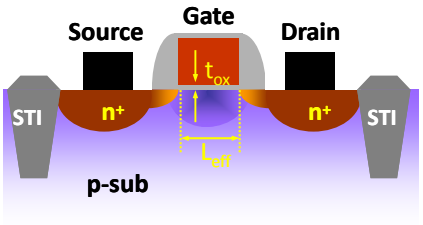


$$I_0 = I_S\left(e^{\frac{V}{V_T}} - 1\right)$$

pn Junction Current Law

Weak Inversion Leakage Current (I)

2 *Static Power Consumption*



As CMOS technology scales down, transistors' t_{ox} & L are reduced. Thus, a lower power supply voltage V_{DD} is used for the reduction of the internal electric fields.

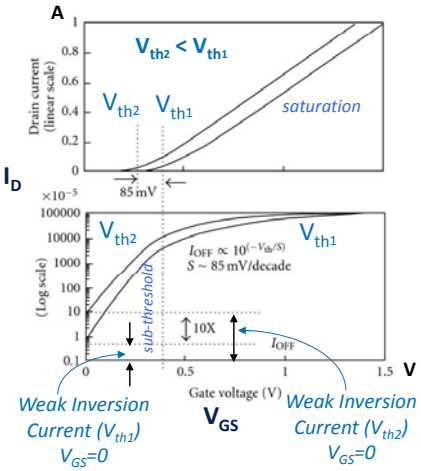
Consequently, also the threshold voltage V_{th} must be reduced to maintain performance, which in turn results in the increment of I_{off} .

Subthreshold slope factor:

$$S = \frac{dV_{GS}}{d(\log(I_D))} = \eta\left(\frac{kT}{q}\right)\ln(10) = 2.3\eta V_T$$

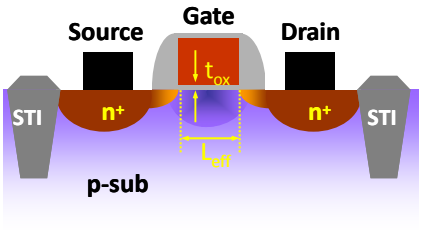
in mV/decade

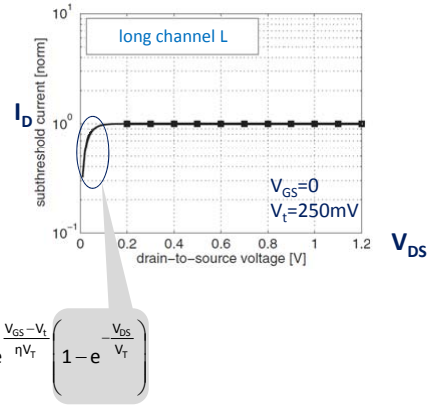
$$I_{off} = \mu C_{ox}\frac{W_{eff}}{L_{eff}}(\eta - 1)V_T^2 e^{\frac{V_{GS}-V_t}{\eta V_T}}\left(1 - e^{-\frac{V_{DS}}{V_T}}\right)$$



Weak Inversion Leakage Current (II)

Static Power Consumption





$$I_{off} = \underbrace{\mu C_{ox}\frac{W_{eff}}{L_{eff}}(\eta-1)V_T^2}_{I_0} e^{\frac{V_{GS}-V_t}{\eta V_T}}\left(1-e^{-\frac{V_{DS}}{V_T}}\right) = I_0 e^{\frac{V_{GS}-V_t}{\eta V_T}}\left(1-e^{-\frac{V_{DS}}{V_T}}\right)$$

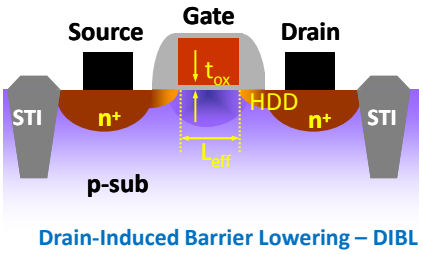
For $V_{DS} > 4\ V_T$ this factor is equal to one!
(where V_T is the thermal voltage (26mV))

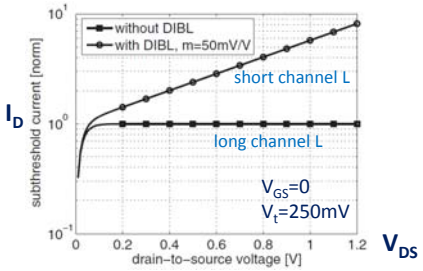
where $\eta = 1+\frac{C_{dm}}{C_{ox}}$ and C_{dm} is the bulk depletion layer capacitance at threshold.

Note that V_t also strongly depends on the temperature (in the range of 0.7 – 1.0 mV/K). The V_t decreases by 70 – 100mV when the temperature increases from 25°C to 125°C.

DIBL Leakage Current

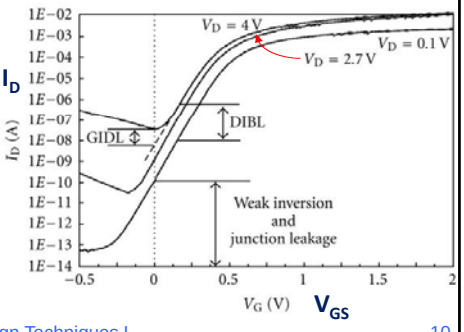
3 *Static Power Consumption*





Drain-Induced Barrier Lowering – DIBL

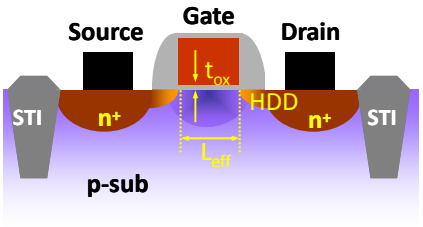
DIBL leakage current appears when the drain potential increases in short L devices. The pn-junction between the drain and the substrate becomes more reverse-biased, so the depletion layer grows and reduces the volume controlled by the gate. Thus, V_{th} decreases as V_{DS} increases: $V_{th}=V_{th,0} - mV_{DS}$

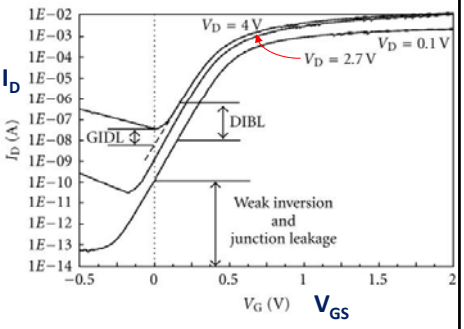


$$I_{off} = I_0 e^{\frac{V_{GS}-V_{th,0}+mV_{DS}}{\eta V_T}} \quad \text{for } V_{DS} > 4\ V_T$$

GIDL Leakage Current

4 *Static Power Consumption*





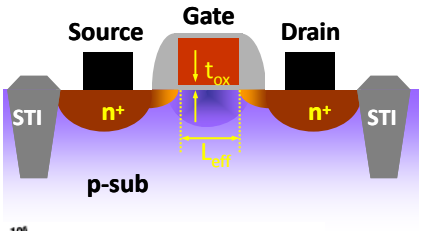
Gate-Induced Drain Leakage – GIDL

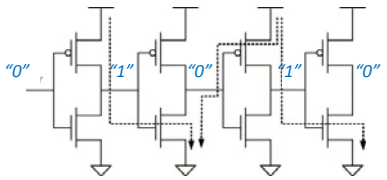
GIDL leakage current is due to the high electric filed, induced by the gate potential, in the gate-drain overlap region.

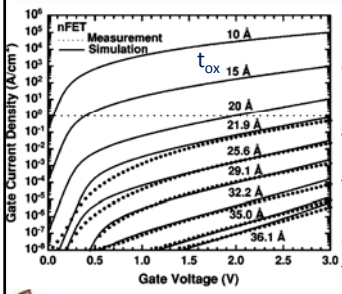
In low-leakage transistors (high V_t and thick gate oxide) the GIDL current dominates the I_{off}.

Gate Oxide Tunneling Leakage Current

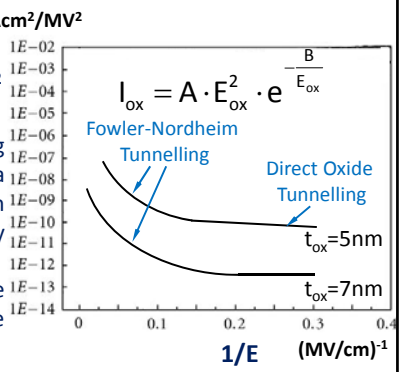
5 *Static Power Consumption*







Gate oxide tunneling leakage current is a great concern in nanometer technology transistors (t_{ox} < 2nm). High-k dielectrics are exploited to alleviate the problem.



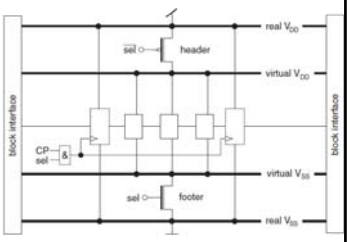
Static Power Consumption Reduction

Power Gating

In power gating we use of a switch device (sleep transistor) between the circuit block and the power supply (V_{DD}, Gnd), aiming to reduce the static power during the idle system states (sleep states).

Trade-offs arise between:

- The maximum delay degradation.
- The requested leakage reduction ratio (LRR).
- The minimum effective power down time (T_{min}).
- The cost.



Design constraints:

- The area overhead, which affects cost and yield.
- The maximum acceptable delay overhead.
- The standby-power specifications that must consider the standby-time of the system.
- The switch (sleep) transistor type and the available options for the standard cells.

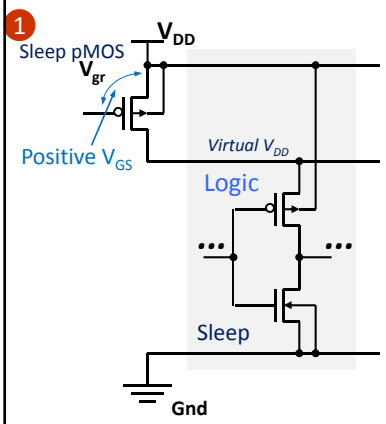
Properties of Power Gating

- High leakage reduction ratio.
- Small impact on performance in active mode.
- Small area overhead due to the: switches, control logic, drivers …
- Independent power gating for different circuit blocks - cores.
- Fast sleep-in and sleep-out periods with small energy overhead.
- Multiple sleep states.

Side Effects:

- The influence of surrounding active blocks by the fast activation of a large block from its sleep mode. Aiming to alleviate this, usually a latency is introduced which increases T_{min}.
- The need to retain the internal states during the sleep mode.
- Block interfaces are required between active and inactive blocks.
- Idle time statistics must be exploited to predict block activation and proper control logic must be implemented for the activation task. This will increase power and consequently T_{min}.

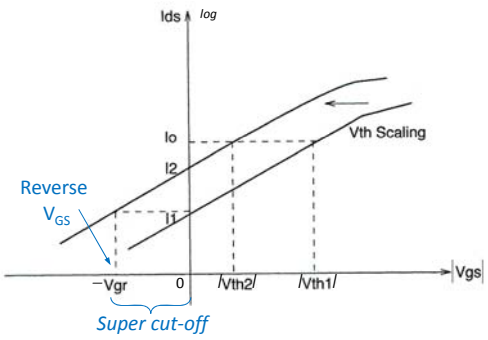
Power Gating with a Reverse V_{GS} Switch



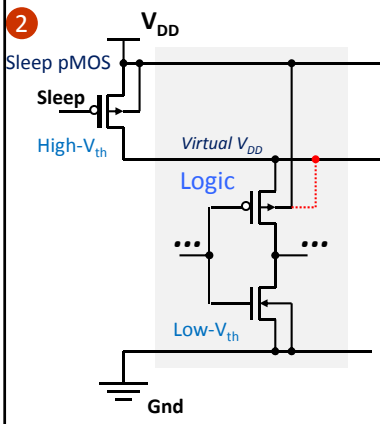
- In active mode a gate voltage V_G lower than 0 is used (boosted gate) aiming to maintain performance.

$$I_{off} = \mu C_{ox} \frac{W_{eff}}{L_{eff}} (\eta - 1) V_T^2 e^{\frac{V_{SG} - |V_{thp}|}{\eta V_T}} \left(1 - e^{-\frac{V_{SD}}{V_T}}\right)$$

- In sleep mode a gate voltage V_G higher than V_{DD} is used for I_{off} reduction.



Power Gating with Multiple V_{th} CMOS



Multiple threshold voltage CMOS (MTCMOS) technology

$$I_{off} = \mu C_{ox} \frac{W_{eff}}{L_{eff}} (\eta - 1) V_T^2 e^{\frac{V_{SG} - |V_{thp}|}{\eta V_T}} \left(1 - e^{-\frac{V_{SD}}{V_T}}\right)$$

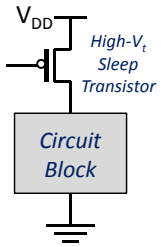
A transistor with higher threshold voltage (e.g. a double gate oxide transistor) is used between the logic circuit and the power supply. In the idle mode of operation this transistor is switched to the cut-off region (sleep mode) by exploiting a dedicated signal (*Sleep*).

Thus, leakage currents due to weak inversion and gate oxide tunneling can be exponentially decreased.

$$V_{th} = V_{th0} \pm \gamma\left(\sqrt{|(-2)\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right)$$

$$\gamma = \frac{t_{ox}}{\varepsilon_{ox}} \sqrt{2q\varepsilon_{Si}N}$$

Threshold Voltage Increase vs Area Overhead



A high-V_t transistor is desirable due to the exponential impact of V_t on subthreshold currents. However, the on-resistance of the transistor will be larger.

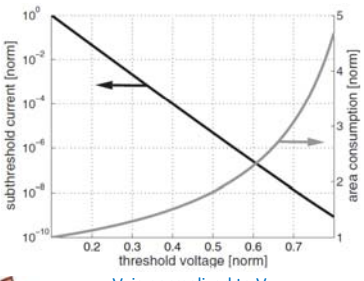
The latter can be compensated by a larger transistor width W, which in turn will increase area overhead and also will increase linearly the subthreshold current. The selection of W with respect to V_t is crucial!

For a desired on-current I_D it stands that:

$$I_D = \mu C_{ox} \frac{W}{L}\left[(V_{DD} - V_t)V_{DSmax} - \frac{V_{DSmax}^2}{2}\right]$$

where $|V_{GS}|=V_{DD}$ and $|V_{DS}|=V_{DSmax}$ which is the max V_{DS} that guarantees the required switching speed of the logic.

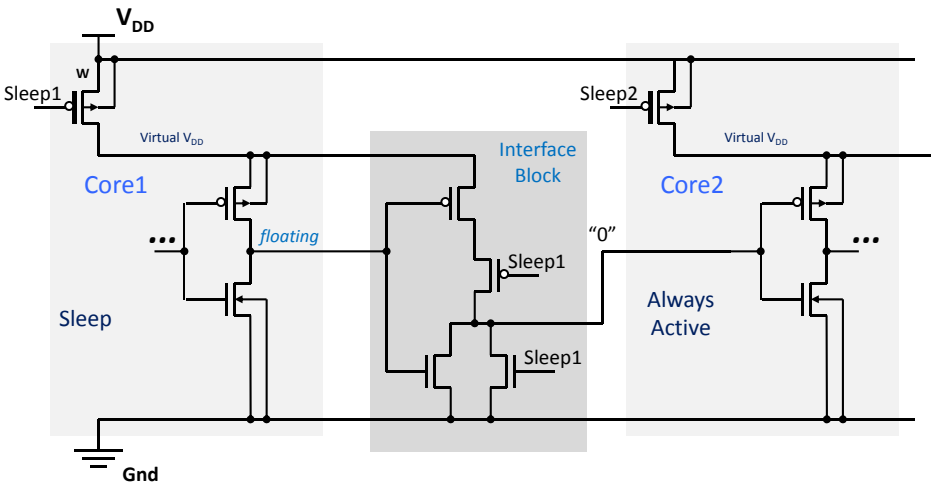
$$W(V_t) = \frac{I_D L}{\mu C_{ox}}\left[(V_{DD} - V_t)V_{DSmax} - \frac{V_{DSmax}^2}{2}\right]^{-1}$$



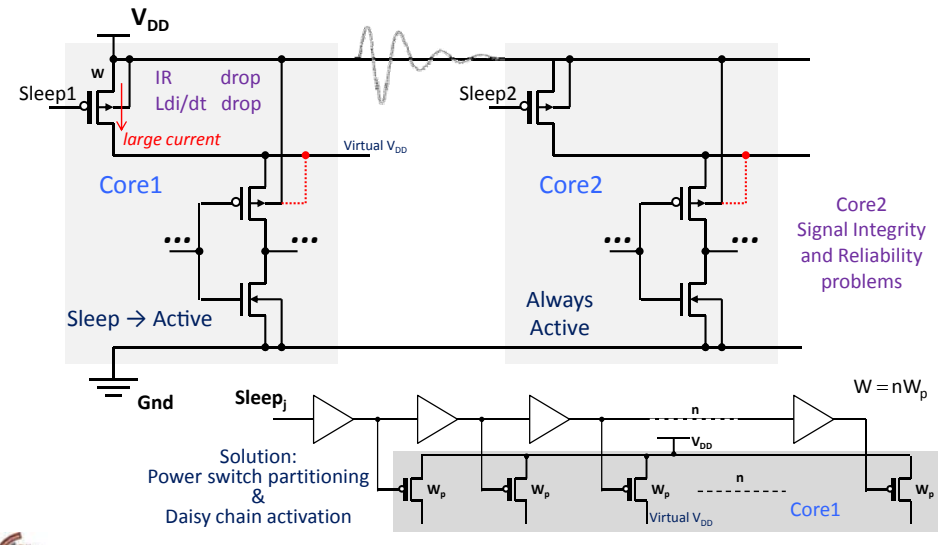
V_t is normalized to V_{DD}

Inserting this W in the I_{off} equation the side graph is drawn. The area overhead grows rapidly while the subthreshold current decreases monotonously!

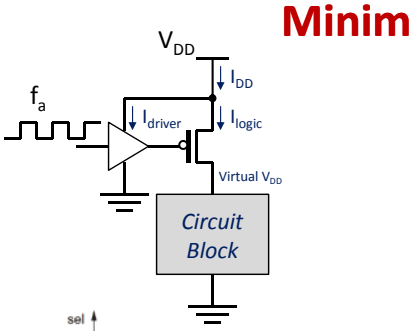
Interfaces in Power Gating

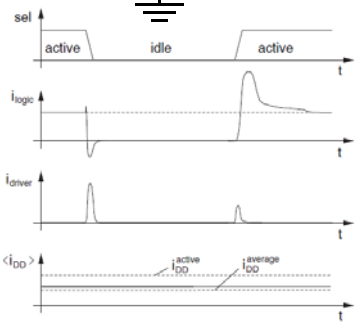


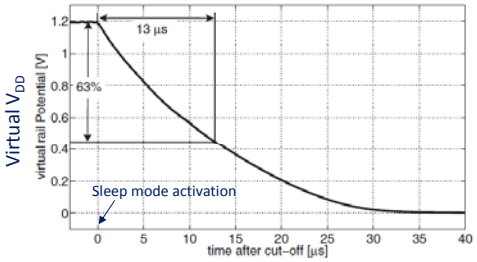
Power Supply Disturbance in Power Gating



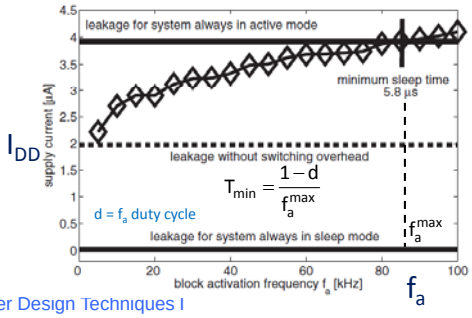
Minimum Idle Time





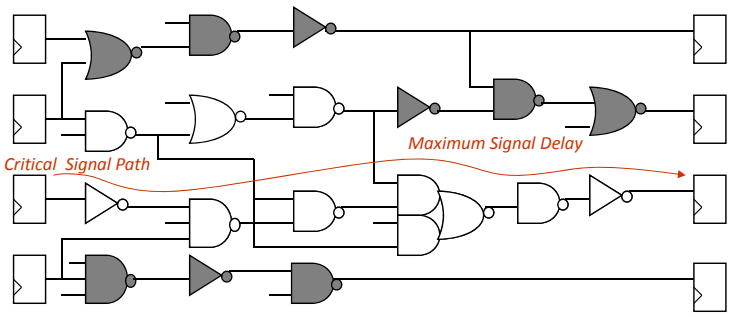


$$\langle I_{DD} \rangle = d \cdot I_{active\_leakage} + (1-d) \cdot I_{idle\_leakage}$$



Multi-Threshold Voltage V_{th} Logic Design

3



The shaded logic blocks have been designed using transistors with higher than the technology nominal threshold voltage in order to reduce the sub-threshold leakage current (weak-inversion leakage current).
Considering that higher threshold voltage can be achieved using thicker gate oxides (insulators), the gate oxide tunneling leakage current can be also reduced.

Active Body Bias Techniques

4

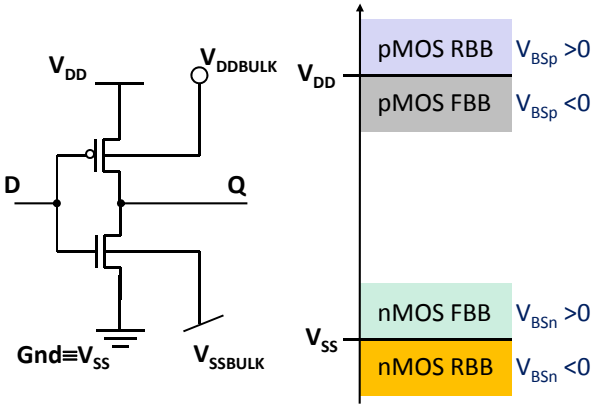
Tunable Threshold Voltage – V_{th}

- Reverse Body Bias – RBB:
  - the absolute threshold voltage increases.
  - the subthreshold leakage current decreses.
  - the performance decreases.

$$I_{off} \approx I_0 e^{\frac{V_{GS}-V_t}{\eta V_T}}$$

$$I_D \approx (V_{GS} - V_t)^{\alpha}$$

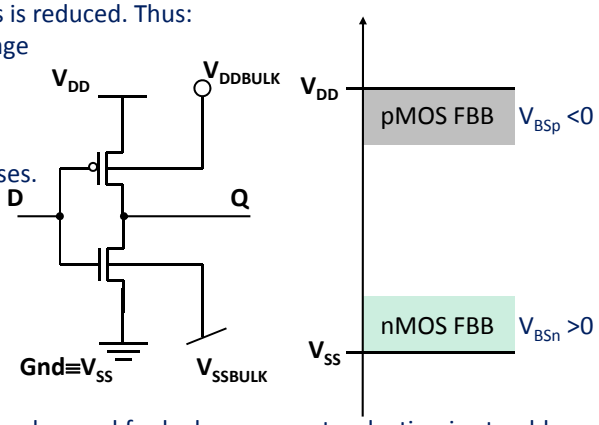
- Forward Body Bias – FBB:
  - the absolute threshold voltage decreases.
  - the subthreshold leakage current increases.
  - the performance increases.



$$V_{th} = V_{th0} \pm \gamma\left(\sqrt{|(-2)\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right)$$

Forward Body Bias

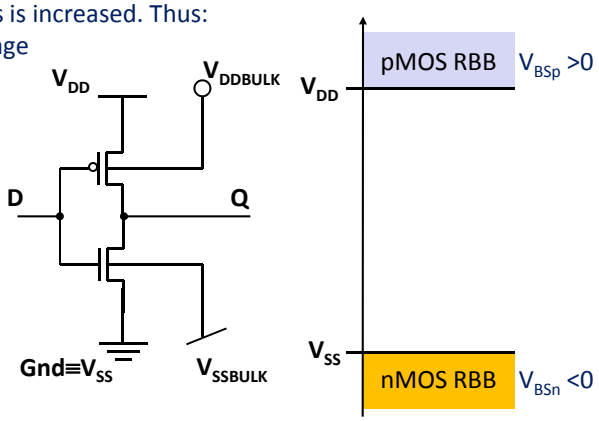
- In Forward Body Bias (FBB) the width of the bulk depletion layer and the drain/source depletion regions is reduced. Thus:
  - the absolute threshold voltage decreases.
  - the subthreshold leakage current increases.
  - the sturation current increases.
  - the drain/source junction capacitances increase.
  - short channel effects (e.g. DIBL) are decreased.
  - the sensitivity to parameter variations is decreased.



- Transistors with high-V_{th} can be used for leakage current reduction in stand-by mode while in active high-performance mode FBB reduces V_{th} and enables fast operation.
- The increased performance can be used to lower V_{DD} and reduce dynamic power.

Reverse Body Bias

- In Reverse Body Bias (RBB) the width of the bulk depletion layer and the drain/source depletion regions is increased. Thus:
  - the absolute threshold voltage increases.
  - the subthreshold leakage current decreases.
  - the sturation current decreases.
  - the drain/source junction capacitances decreases.
  - the sensitivity to parameter variations is increased.



- Transistors with nominal V_{th} can be used for high-performance operation while in stand-by mode RBB increases V_{th} and reduces static power.

Variable Threshold Voltage V_{th}

Reverse Body Bias – RBB

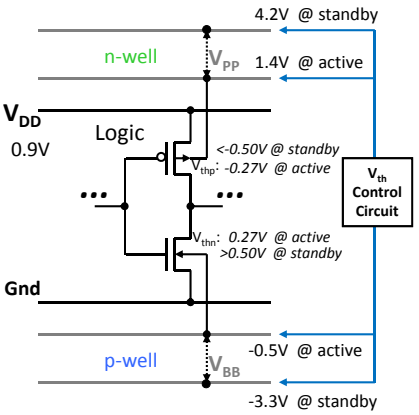
Substrate (body) voltage adaptation has been proposed, in order to increase the transistor threshold voltage in the idle state. In this state, the substrate of the pMOS transistors is biased to higher than V_{DD} voltage levels while the substrate of nMOS transistors is biased to negative voltage levels.

Thus, the weak inversion current is exponentially decreased, according to the pertinent expression, since the absolute threshold voltage V_{th} is increased.

Moreover, the DIBL current is also decreased.

$$V_{th} = V_{th0} \pm \gamma\left(\sqrt{|(-2)\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right)$$

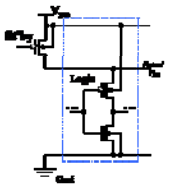
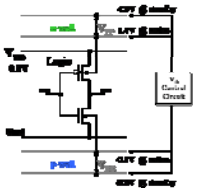
$$\gamma = \frac{t_{ox}}{\varepsilon_{ox}}\sqrt{2q\varepsilon_{Si}N}$$



Variable threshold voltage – VTCMOS with Reverse Body (substrate) Bias – RBB

Kuroda et.al., IEEE J. Solid-State Circuits, 31(11), 1996

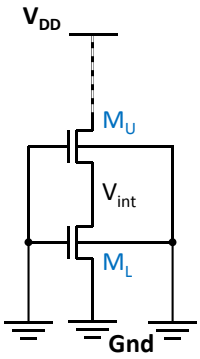
Advantages – Disadvantages

	Multi-threshold voltages MTCMOS	Variable threshold voltages VTCMOS
Technique		
Characteristics	*+ Applicability* *+ Typical wells* *- Stacked MOS transistors* *- I_{DDQ} testing* *- Dedicated Flip-Flop design* *- No V_{th} variation adaptation*	*+ No stacked MOS (speed)* *+ I_{DDQ} testing* *+ V_{th} variation adaptation* *- Silicon area cost* *- Triple-well technologies are required* *- pn-junction leakage current is increased* *-Body bias generator is required* *- Reduced applicability with technology scaling*

180nm → $|V_{BS}|$=500mV ⇒ μείωση I_{off} 4-5×
130nm → $|V_{BS}|$=500mV ⇒ μείωση I_{off} 3-3.5×

The Transistor Stacks Technique

5



Case of two nMOS transistors connected in series to form a stack.

In general: $V_{th} = V_{th0} - mV_{DS} - gV_{BS}$

where m the DIBL coefficient and gV_{BS} a linear approximation of the body effect.

For M_L it stands that V_{GSL}=0 while no body effect exists (V_{BSL}=0). Moreover, V_{DSL}=V_{int} and V_{int}<V_{DD}. Thus, the DIBL current is reduced and the I_{off} is:

$$I_{offL} = I_0 e^{\frac{-V_{t0}+mV_{int}}{\eta V_T}} \left(1 - e^{-\frac{V_{int}}{V_T}}\right)$$

For M_U it stands that V_{GSU}=−V_{int} while reverse body bias exists (V_{BSU}=−V_{int}). Moreover, V_{DSU}=V_{DD}−V_{int}>0. Thus, the I_{off} is:

$$I_{offU} = I_0 e^{\frac{-V_{int}-V_{t0}+m(V_{DD}-V_{int})+gV_{int}}{\eta V_T}} \left(1 - e^{-\frac{(V_{DD}-V_{int})}{V_T}}\right)$$

It stands that: I_{offU} = I_{offL}

References

- "*Low Power Design Methodologies*," J. Rabaey and M. Pedram, Springer, 1997.
- "*Power Management of Digital Circuits in Deep Sub-Micron CMOS Technologies*," S. Henzler, Springer, 2006.
- "*Low-Power Digital VLSI Design*," A. Bellaouar and M. Elmasry, Kluwer Academic Publishers, 1996.
- "*A 1.5V Full-Swing Bootstrapped CMOS Large Capacitive-Load Driver Circuit Suitable for Low-Voltage CMOS VLSI*," J. Lou and J. Kuo, IEEE Journal of Solid-State Circuits, vol. 32, no. 1, pp. 119-121, 1997.
- "*A 0.9-V, 150-MHz, 10-mW, 4 mm , 2-D Discrete Cosine Transform Core Processor with Variable Threshold-Voltage (VT) Scheme*," T. Kuroda et.al., IEEE Journal of Solid-State Circuits, vol. 31, no. 11, pp. 1770-1779, 1996.
- "*Low-Power CMOS Digital Design*," A. Chandrakasan, S. Sheng and R. Brodersen, IEEE Journal of Solid-State Circuits, vol. 27, no. 4, pp. 473-484, 1992.
- "*Minimazing Power Consumption in Digital CMOS Circuits*," A. Chandrakasan, R. Brodersen, Proceedings of the IEEE, vol. 83, no. 4, pp. 498-523, 1995.