EECS 427
VLSI Design

Lecture 11: Dynamic Logic Families
Prof. Pinaki Mazumder
Winter 2013

Adapted from Harris, Rabaey, Blaauw, Zhang, Sylvester, and others

Outline

• Basic domino gate
• Issues in dynamic gates
• Domino cascading
• Footless domino
• NORA/Zipper logic
• Multiple-output domino logic
• Compound domino
• Dual-rail domino
• Self-reseting domino
• Limited Switch dynamic logic
On Optimal Tapering of FET Chains in High-Speed CMOS Circuits

Li Ding, Student Member, IEEE, and Pinaki Mazumder, Fellow, IEEE

Domino / Static $C_{in}/C_{out}$

3-input OR gate

Static:
$C_i = 7$
$C_n = 9$

Dynamic:
$C_i = 2$
$C_n = 7$

Skew for eval
Basic Domino Gate

• Divide the clock in 2 phases:
  - Precharge
    • Output Low
    • Dyn. Cap precharge
    • PDN Off
  - Evaluate
    • Conditional discharge
    • Input must be stable and monotonic L→H

Advantages:
- Faster than CMOS
- Input capacitance is lower
- Early switch point
- Inverter P/N > 2 (only rising delay important)
**Basic Domino Gate**

- **Disadvantages:**
  - Low noise margin
  - Charge sharing
  - Leakage currents
  - Internal capacitance charge sensitive to noise

**Why Domino?**

Like falling dominos!
Footless Domino

Precharge after inputs precharge
Evaluate before inputs valid

Inverter
NAND2
NOR2

Unfooled
\[ g_d = \frac{1}{3}, \quad p_d = \frac{2}{3} \]
\[ g_d = \frac{2}{3}, \quad p_d = 3/3 \]

Footed
\[ g_d = \frac{2}{3}, \quad p_d = 3/3 \]
\[ g_d = 3/3, \quad p_d = 4/3 \]

*FIGURE 9.25 Catalog of dynamic gates*
**Dynamic CMOS**

- In static circuits at every point in time (except when switching) the output is connected to either GND or V\textsubscript{DD} via a low resistance path.
  - fan-in of \( n \) requires \( 2n \) (\( n \) N-type + \( n \) P-type) devices

- Dynamic circuits rely on the temporary storage of signal values on the capacitance of high impedance nodes.
  - requires on \( n + 2 \) (\( n \)+1 N-type + 1 P-type) transistors

**Dynamic Gate**

Two phase operation

- Precharge (CLK = 0)
- Evaluate (CLK = 1)
Dynamic Gate

Two phase operation
- Precharge (Clk = 0)
- Evaluate (Clk = 1)

Conditions on Output

- Once the output of a dynamic gate is discharged, it cannot be charged again until the next precharge operation.
- Inputs to the gate can make at most one transition during evaluation.
- Output can be in the high impedance state during and after evaluation (PDN off), state is stored on $C_L$. 
Properties of Dynamic Gates

- Logic function is implemented by the PDN only
  - number of transistors is $N + 2$ (versus $2N$ for static complementary CMOS)
- Full swing outputs ($V_{OL} = GND$ and $V_{OH} = V_{DD}$)
- Non-ratioed - sizing of the devices does not affect the logic levels
- Faster switching speeds
  - reduced load capacitance due to lower input capacitance ($C_{in}$)
  - reduced load capacitance due to smaller output loading ($C_{out}$)
  - no $I_{sc}$, so all the current provided by PDN goes into discharging $C_L$

Properties of Dynamic Gates

- Overall power dissipation usually higher than static CMOS
  - no static current path ever exists between $V_{DD}$ and GND (including $P_{sc}$)
  - no glitching
  - higher transition probabilities
  - extra load on Clk
- PDN starts to work as soon as the input signals exceed $V_{Tn}$, so $V_M$, $V_{IH}$ and $V_{IL}$ equal to $V_{Tn}$
  - low noise margin ($NML$)
- Needs a precharge/evaluate clock
Issues in Dynamic Design 2: Charge Sharing

Charge stored originally on $C_L$ is redistributed (shared) over $C_L$ and $C_A$ leading to reduced robustness.

Charge Sharing Example

Charge stored originally on $C_L$ is redistributed (shared) over $C_L$ and $C_A$ leading to reduced robustness.
Charge Sharing

case 1) if $\Delta V_{out} < V_{Tn}$

$$C_L V_{DD} = C_L V_{out(t)} + C_a (V_{DD} - V_{Tn}(V_X))$$

or

$$\Delta V_{out} = V_{out(t)} - V_{DD} = \frac{C_a}{C_L} (V_{DD} - V_{Tn}(V_X))$$

case 2) if $\Delta V_{out} > V_{Tn}$

$$\Delta V_{out} = -V_{DD} \left( \frac{C_a}{C_a + C_L} \right)$$

Solution to Charge Redistribution

Precharge internal nodes using a clock-driven transistor (at the cost of increased area and power)
Issues in Dynamic Design 3: Backgate Coupling

Backgate Coupling Effect

© Digital Integrated Circuits 2nd
Combinational Circuits
Issues in Dynamic Design 4: Clock Feedthrough

Coupling between Out and Clk input of the precharge device due to the gate to drain capacitance. So voltage of Out can rise above $V_{DD}$. The fast rising (and falling edges) of the clock couple to Out.

Clock Feedthrough

Clock feedthrough
Issues: Leakage

- Dynamic node is floating during evaluation
  - Leakage current of NMOS can discharge it

Half latch devices:
- Size PMOS to replenish the leakage current
- Limits width OR gates
Domino - with and without keeper

Alternative to half latch device

- Eliminates fight
- But only a few PMOS in series
  - More input cap
- OPL?
Issues: Charge sharing

- In evaluate, dynamic node charge is shared with internal node caps
- Node was discharged in previous cycle

Issues: Charge sharing

- Issues with precharging internal node:
  - PDN becomes slower (more internal cap.)
  - Higher voltage to discharge

Alternative: NMOS

Lower voltage discharge

- More Clock Load
- Higher speed
- Same Charge
- Sharing Protection
Issues: Timing

• Delay is dependent on the timing of inputs

First gate in domino chain ↑ delay

On Circuit Techniques to Improve Noise Immunity of CMOS Dynamic Logic

Li Ding, Member, IEEE, and Pinaki Mazumder, Fellow, IEEE

<table>
<thead>
<tr>
<th>Class</th>
<th>Technique</th>
<th>Reference</th>
<th>Illustration</th>
<th>Num. tran.</th>
<th>Input load a</th>
<th>Clock load b</th>
<th>Driving path c</th>
<th>DC current d</th>
<th>All noises e</th>
<th>All noise f</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Always-on keeper</td>
<td>[3]</td>
<td>Fig. 6(a)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Feedback keeper</td>
<td>[4]</td>
<td>Fig. 6(b)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Feedback keeper</td>
<td>[4]</td>
<td>Fig. 6(b)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Conditional feedback</td>
<td>[5]</td>
<td>Fig. 6(c)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Precharge internal</td>
<td>[9]</td>
<td>Fig. 3(a)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Precharge</td>
<td>[10]</td>
<td>Fig. 3(b)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>C</td>
<td>PMOS pull-up</td>
<td>[11]</td>
<td>Fig. 4(a)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>NMOS pull-up (feedback)</td>
<td>[12]</td>
<td>Fig. 4(b)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Mirror technique</td>
<td>[13]</td>
<td>Fig. 4(c)</td>
<td>N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Twin transistor</td>
<td>[15]</td>
<td>Fig. 4(d)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Complementary p-network</td>
<td>[17]</td>
<td>Fig. 5(a)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>CMOS inverter</td>
<td>[19]</td>
<td>Fig. 5(b)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Gated CMOS inverter</td>
<td>[20]</td>
<td>Fig. 5(c)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>Triple transistor</td>
<td>[21]</td>
<td>Fig. 5(d)</td>
<td>2N</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>

For improved noise immunity of dynamic logic gates using a keeper, the designer can use:

(a) Weak always-on keeper [2]  
(b) Feedback keeper [4]  
(c) HS feedback keeper [5]  
(d) Complementary feedback keeper [23]

*Symbol o represents “good” and symbol x represents “not good.”
Domino cascading

- Input must be stable during eval.

Footless Domino

- First stage has footer.
- Footer is not needed for other gates
  - Inputs must precharge low before dynamic node precharge
  - Delay clocks $\Phi_1, \Phi_2$
Footless Domino

---

**Advantages:**
- Faster than classic domino → one less NMOS
  - Why is footer necessary in the first stage?
  - Input not guaranteed during precharge

---

**Disadvantages:**
- Use of different clocks → Can not simply delay the clock (fast inputs during eval)
- Reduced precharge time for later stages
  - Tradeoff of sizing up PMOS (increase dynamic cap.) vs 1 less NMOS in PDN (footless)
  - But can use a footed stage after a footless stage to recover
• Cascade basic dynamic gates with different evaluation networks (PDN, PUN)

- Advantages:
  - Eliminates the inverter delay (but no drive for long interconnect)
  - Fast
NORA/Zipper Logic

• Disadvantages:
  - Not very good for large output loads (Does not use an output inverter)
  - Big and slow PMOS PUN
  - Bad noise margins → Can add keeper
  - Noise on dynamic node
  - Cannot make arbitrary connections

Dual-Rail Domino

• Only non-inverting gates in domino
  - Dual-rail is required for general logic functions
  - Double the number of transistors
• If evaluation trees are shared, internal nodes are precharged for any input pattern

• Not completely safe. Why?
  - No keeper set. On when inputs have not arrived
  - No fight
Multiple-Output Domino (MODL)

- Implement more logic per domino stage
- Slowdown top output but more work done
- Very common

Compound Domino

- Add static gates to evaluate logic
- Need to add half latches to every output node
Compound Domino

- Reduce the number of transistors in a stack (faster)
  - Due to $V_{dd}$ scaling no more than 4 transistors in a stack ($V_{dd} \sim 4 \times V_{dsat}$)
  - But use two PMOS in NOR gate to drive output high

- What static gate can you add?

Out = $AB \cdot \overline{CD}$
Out = $AB + CD$
Self Resetting Domino

Non clocked precharge. Precharge controlled by output

Dynamic node floating
Self Resetting Domino

NO FIGHT
PMOS OFF

Self Resetting Domino
Self Resetting Domino

Input must be precharged before pulse ripples through

Self Resetting Domino
Self Resetting Domino

Timing constraints:
- $f$ drops after the inputs go low
- $f$ pulls up before the next input
- Multiple inputs must line up in time

Input must happen after precharge ripples through
Self Resetting Domino

- Advantages:
  - No clock
  - Fast eval because no footer
  - More time for pre-charge than standard footless domino

- Disadvantages:
  - Timing constraints
  - Stability after precharge
  - Sensitive to process variations
  - Very difficult in practice → pulse everything
Issues: Miller capacitance

More charge to pull-down by NMOS

Only affects precharge

Sample Set Differential Logic (SSDL)
Sample Set Differential Logic (SSDL)

\[ \Phi = 0 \text{ Sample} \]

\[ V_{\text{true}} \quad 0 \quad V_{\text{comp}} \]

Differential Tree

\[ \Phi = 1 \text{ Amplify} \]

\[ V_{\text{true}} \quad 1 \quad V_{\text{comp}} \]

Differential Tree
Sample Set Differential Logic (SSDL)

- Advantages:
  - Fast

- Disadvantages:
  - Short circuit power
  - Every gate in one clock phase

Question #1: Why do we use static inverter in domino logic? Or, why don’t we cascade two dynamic logic gates?

Prof. Pinaki Mazumder at University of Michigan
Question #2: What is the difference between Clocked CVSL and SSDL?

Sample Set Differential Logic (SSDL)

Prof. Pinaki Mazumder at University of Michigan

Question #3: Why do we need two clocks or delayed clocks for Footless Domino?
Lecture 12 - Static Power

WH 5.3
Adapted from Weste & Harris, and Rabaey & Chandrakasan
Prof. Pinaki Mazumder

Topics

• Leakage mechanisms
  – Subthreshold leakage
  – Gate oxide leakage

• Leakage reduction methods
  – State assignment
  – MTCMOS
  – Dual-Vth design
  – VTCMOS
Leakage mechanisms

• **I1: Reverse-bias p-n junction**
  - Reverse-biased p-n junctions current: \( I_D = I_S \left( \frac{V_{DS}}{V_T} \right) e^{\frac{V_{DS}}{V_T}} - 1 \)
  - Typically < 1 fA/mm² (negligible)
  - Depends on area and perimeter of diffusion regions
  - Also: Band to Band tunneling (BTBT)

Leakage mechanisms

• **I2: Weak inversion or subthreshold leakage current**
  - Increased voltage increases drain depletion extending to the source → lowers the potential barrier
  - Dominant effect in modern devices
Sub-Threshold Conduction

\[ I_D = I_0 e^{\frac{qV_{GS}}{nkT}} \left( 1 - e^{-\frac{qV_{DS}}{nkT}} \right) \]

**The Slope Factor**

\[ I_D \sim I_0 e^{\frac{qV_{GS}}{nkT}}, \quad n = 1 + \frac{C_D}{C_{ox}} \]

\[ S = n \left( \frac{kT}{q} \right) \ln(10) \]

Typical values for S:

- 60..100 mV/decade

---

Subthreshold Leakage

- Subthreshold leakage exponential with \( V_{gs} \)

\[ I_{ds} = I_{ds0} e^{\frac{V_{gs} - V_{th} - qV_{th} - kT}{n\nu_T}} \left( 1 - e^{\frac{-V_{th}}{n\nu_T}} \right) \]

- \( n \) is process dependent – typically 1.3-1.7
- \( \nu_T = kT/q \)
- threshold voltage: \( V_{th} \)

\[ I_{ds} = I_{ds0} 10^\frac{V_{gs} + qV_{th} - V_{th} - kT}{S} \left( 1 - e^{\frac{-V_{th}}{\nu_T}} \right) S = \left[ \frac{d\log_{10} I_{ds}}{dV_{gs}} \right]^{-1} = n\nu_T \ln 10 \]

- \( S \approx 100 \) mV/decade @ room temperature

The threshold voltage decreases with increasing \( V_{ds} \). This effect, called the drain-induced barrier lowering, or DIBL, causes the threshold potential to be a function of \( V_{ds} \)
Threshold Voltage

\[ V_T = V_{T0} + \gamma (\sqrt{2\Phi_F} - V_{SB}) - \sqrt{2\Phi_F} \]

\[ = V_{T0} - K_Y V_{BS} = -\eta V_{GS} - K_Y V_{BS} \]

\( V_{BS} \) (V)

\( \gamma = \frac{V_{GS}}{V_{PD}} \)

Linear

Approximation

Model

\[ I_{ds} = I_{off} \ 10^{-\frac{V_{gs} + \eta(V_{dd} - V_{dd})}{2}} \left( 1 - e^{-\frac{V_{dd}}{V_{T}}} \right) \]

Drain-Induced Barrier Lowering (DIBL)

\[ I_{ds} = I_{off} \ 10^{-\frac{V_{gs} + \eta(V_{dd} - V_{dd})}{2}} \left( 1 - e^{-\frac{V_{dd}}{V_{T}}} \right) \]

- Electric field from drain affects channel

- More pronounced in small transistors where drain to channel coupling is stronger

- Drain-Induced Barrier Lowering effectively reduces threshold voltage

- High drain voltage causes leakage to increase.
**Body Coefficient / Vds Dependence**

For NMOS: lower body voltage relative to source voltage (reverse bias)
- Increases effective Vth
- Reduces leakage

Vds dependence
- For $V_{ds} > 4V_{T}$ leakage current independent of Vds (other than DIBL)
- For $V_{ds} < 2V_{T}$ leakage current drops rapidly with lower Vds

**Leakage Mechanisms**

I3: GIDL - Gate Induced Drain Leakage
- Negative gate / Positive drain
- Thins out drain depletion causing drain to well leakage near gate
- Generates a tunneling current
Subthreshold Leakage Roundup

- For $V_{ds} > 50$ mV
  
  \[
  I_{sub} \approx I_{off} \frac{V_{gs} + \eta(V_{ds} - V_{DD}) - k_s V_{sb}}{S}
  \]

- $I_{off} =$ leakage at $V_{gs} = 0, V_{ds} = V_{DD}$

Typical values in 65 nm

- $I_{off} = 100$ nA/µm @ $V_{th} = 0.3$ V
- $I_{off} = 10$ nA/µm @ $V_{th} = 0.4$ V
- $I_{off} = 1$ nA/µm @ $V_{th} = 0.5$ V

DIBL coefficient: $\eta = 0.1$
Body effect coefficient: $k_s = 0.1$
S = 100 mV/decade
Leakage Mechanisms

- **I4: Gate Oxide tunneling**
  - Thinner oxides cause an increase tunneling
  - Highly dependent on oxide material and thickness
Gate Leakage

- Carriers tunnel through very thin gate oxides
- Exponentially sensitive to $t_{ox}$ and $V_{DD}$

$$I_{gate} = WA \left( \frac{V_{DD}}{t_{ox}} \right)^2 e^{-B \frac{t_{ox}}{V_{DD}}}$$

- $A$ and $B$ are tech constants
- Greater for electrons
  - So NMOS gates leak more

- Negligible for older processes ($t_{ox} > 30 \text{ Å}$)
- Critically important at 65 nm and below ($t_{ox} \approx 12 \text{ Å}$)
  - But: improved again with High-K metal gate transistors

Fundamental Leakage Levers

- Increase $V_{th}$: ~10x leakage reduction for every 100mV
  - But: bad for delay
  $$\tau \propto \frac{V_{dd}}{(V_{dd} - V_t)^\alpha}$$

- Reduce temperature: ~5.2X reduction / 10 degree C

- Stacking transistors
Stack Effect

- Series OFF transistors significantly reduce leakage
  - $V_x > 0$, so $N_2$ has negative $V_{gs}$

$$I_{sub} = I_{off} \frac{\eta(V_x - V_{DD})}{S} \approx I_{off} \frac{-(V_x + \eta(V_{DD} - V_{TH}) - k_x)}{S}$$

$$V_x = \frac{\eta V_{DD}}{1 + 2\eta + k_x}$$

$$I_{sub} = I_{off} \frac{\eta V_{DD}}{S} \approx I_{off} \frac{-\eta V_{DD}}{S}$$

- Leakage through 2-stack reduces $\sim 10x$

Stacking and Leakage

<table>
<thead>
<tr>
<th># of stack</th>
<th>Leakage current (pA)</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>258</td>
<td>X 1</td>
</tr>
<tr>
<td>2</td>
<td>36.1</td>
<td>X 7.1</td>
</tr>
<tr>
<td>3</td>
<td>19.8</td>
<td>X 13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># of stack</th>
<th>Leakage current (pA)</th>
<th>Reduction</th>
<th>Reduction per HVT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.25</td>
<td>X 206</td>
<td>X 1</td>
</tr>
<tr>
<td>2</td>
<td>0.185</td>
<td>X 1394</td>
<td>X 6.8</td>
</tr>
<tr>
<td>3</td>
<td>0.122</td>
<td>X 2115</td>
<td>X 10.3</td>
</tr>
</tbody>
</table>
State Assignment

- Only a few states have significant leakage
  - Dominant leakage states have only one transistor OFF in any path from \( V_{dd} \) to Gnd

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Leakage Current (pA)</th>
<th>Leaking Transistors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10.537</td>
<td>N1, N2, N3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>18.534</td>
<td>N1, N2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>18.234</td>
<td>N1, N3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>135.772</td>
<td>N1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20.350</td>
<td>N2, N3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>102.672</td>
<td>N2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>100.970</td>
<td>N3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>192.174</td>
<td>P1, P2, P3</td>
</tr>
</tbody>
</table>

Leakage currents in pA. NMOS width = 480nm PMOS width = 320nm

State Dependence of Leakage

- Circuit state is partially unknown in sleep state
- Leakage variation is less for entire circuit than for individual gates

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Leakage Current (nA)</th>
<th>Max / Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adder1</td>
<td>256.8 283.1 309.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Control</td>
<td>33.8 45.97 60.23</td>
<td>1.78</td>
</tr>
<tr>
<td>Decoder</td>
<td>1702.5 1914.3 2122.1</td>
<td>1.25</td>
</tr>
<tr>
<td>Nand4</td>
<td>0.07 0.76 7.1</td>
<td>101.4</td>
</tr>
<tr>
<td>OAI21</td>
<td>0.84 7.73 17.78</td>
<td>21.2</td>
</tr>
<tr>
<td>Tinv</td>
<td>0.37 1.89 5.76</td>
<td>15.6</td>
</tr>
<tr>
<td>AOI21</td>
<td>2.44 8.51 17.23</td>
<td>7.1</td>
</tr>
</tbody>
</table>
State Assignment

• Since the leakage of a logic gate depends on its input, find the input to a combinational circuit that minimizes leakage
  – 30%-40% leakage variation depending on input vector

• Modify latches
  – Sleep signal moves pre-determined values as inputs into combinational circuit

Power Gating – aka MTCMOS

• Turn OFF power to blocks when they are idle
  – Use virtual V_DD and Gnd
  – “Gate” outputs to prevent invalid logic levels at next block
  – Use HVT header/footer

• Voltage drop across sleep transistor during normal operation
  – Size the transistor wide enough to minimize impact

• Switching sleep transistor costs dynamic power
  – Only justified when circuit sleeps long enough
State Retaining MTCMOS Latch

Sleep Transistor Layout
Dual-Thresholds Inside a Logic Block

- Minimum energy consumption is achieved if **all** logic paths are critical (have the same delay)
- Use lower threshold on timing-critical paths
  - Assignment can be done on a per gate or transistor basis; no clustering of logic is needed
  - No level converters needed

Vth Assignment Granularity

- Vth assignment can be performed at different levels of granularity
  - Gate level assignment
  - Pull up network / Pull down network based assignment (half gate)
    - Single Vth in pull up or pull down networks
  - Stack based assignment
    - Single Vth in series connected transistors
  - Individually assignment within transistor stacks
    - Possible area penalty (see right)
- Number of library cells increases with finer control
  - Better leakage / delay trade-off
  - Harder for synthesis tools to handle
Dynamic Body Bias

Active mode
Forward body bias

Idle mode
Reverse body bias
Triple well needed

Variable Threshold CMOS (VTCMOS)
Energy Dissipation at Subthreshold and Superthreshold Voltages

Leakage Reduction Overview

MTCMOS  Dual Threshold  State Assignment  Variable $V_t$

Low $V_t$ Logic

High $V_t$

Variable $V_t$ Logic

Source: [Johnson, et al., DAC99]
## Power and Energy Design Space

<table>
<thead>
<tr>
<th></th>
<th>Constant Throughput/Latency</th>
<th>Variable Throughput/Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>Design Time</td>
<td>Non-active Modules</td>
</tr>
<tr>
<td>Active</td>
<td>Logic Design</td>
<td>Clock Gating</td>
</tr>
<tr>
<td></td>
<td>Sizing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low C circuits</td>
<td></td>
</tr>
<tr>
<td>Leakage</td>
<td>Multi-$V_{th}$</td>
<td>Sleep Transistors</td>
</tr>
<tr>
<td></td>
<td>Stack effect</td>
<td>State assignment</td>
</tr>
<tr>
<td></td>
<td>Variable $V_{th}$</td>
<td>Variable $V_{th}$</td>
</tr>
</tbody>
</table>

## Processor Power Management

- Software power control - power management
  - **DOZE**: Most units stopped except on-chip cache memory (cache coherency)
  - **NAP**: Cache also turned off, PLL still on, time out or external interrupt to resume
  - **SLEEP**: PLL off, external interrupt to resume

Deeper sleep mode consumes less power. Deeper sleep mode requires more latency to resume.

<table>
<thead>
<tr>
<th>Mode</th>
<th>66MHz</th>
<th>80MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>No power mgmt</td>
<td>2.18W</td>
<td>2.54W</td>
</tr>
<tr>
<td>Dynamic power mgmt</td>
<td>1.89W</td>
<td>2.20W</td>
</tr>
<tr>
<td>DOZE</td>
<td>307mW</td>
<td>366mW</td>
</tr>
<tr>
<td>NAP</td>
<td>113mW</td>
<td>135mW</td>
</tr>
<tr>
<td>SLEEP</td>
<td>89mW</td>
<td>105mW</td>
</tr>
<tr>
<td>SLEEP without PLL</td>
<td>18mW</td>
<td>19mW</td>
</tr>
<tr>
<td>SLEEP without clock</td>
<td>2mW</td>
<td>2mW</td>
</tr>
</tbody>
</table>

10 cycles to wake up from SLEEP, 100us to wake up from SLEEP+.
Conclusions

- Lots of recent work on circuit and technology techniques to reduce static power
  - Standby mode leakage reduction can be orders of magnitude, may lose state, takes time to switch in and out of standby mode
  - Active mode leakage reduction is a tougher problem, smaller savings (<50% typically), must be ready for inputs to toggle at any time

Energy Dissipation at Subthreshold and Superthreshold Voltages

\[
\frac{\partial E_T}{\partial V_{DD}} = 2C_{\text{ox}}V_{DD} + 2W_{\text{eff}}L_{DP}K_C V_{DD} e^{-V_{DD}/nV_{th}}
\]

\[
V_{DD_{opt}} = nV_{th}(2 - \text{lambertW}(\beta))
\]

\[
\beta = \frac{-2C_{\text{ox}}}{W_{\text{eff}}L_{DP}K_C} e^2 > e^{-1}
\]

Lambert W function, \( W = \text{lambertW}(x) \)

Solution to the equation \( W e^W = x \)

\[
V_{DD_{opt}} = nV_{th} \ln \left( \frac{fK_C L_{DP} V_{DD_{opt}}}{I_{th}} \right)
\]