Outline

- Motivation
- Problem Statement
- Implementation Overview
- CAD Level Solution
- Experimental Results
- Conclusion
Motivation

caused due to temperature and process variations

S.Borkar [Intel]
Problem Statement

- Determine the right amount of body bias to compensate for
  - process variations
  - temperature variations
- Generalized framework for
  - one-time compensation
  - run-time compensation
- Provide a CAD level solution to the problem of determining the exact body bias values.
Our approach

Temperature Adaptive Body Bias (TABB)

Ideal Process Conditions

Determine voltages $(v_{bn}, v_{bp})_{TABB}$

Process Adaptive Body Bias (PABB)

Ambient Temperature Conditions

Determine voltages $(v_{bn}, v_{bp})_{PABB}$

Process Temperature Adaptive Body Bias (PTABB)

Compute voltages $(v_{bn}, v_{bp})_{PTABB}$

Works at all temperatures for within-die variations
System Level Block Diagram

BBG: Body Bias Generator
(can be central or local)

Critical Path Replica and Phase Detector Circuitry

Circuit Block with Biasable NWELL and PWELL

Temperature Sensor

ROM (Look up Table)

From Body Bias Generator

To Body Bias Generator

V_{bp}  V_{bn}
Independence of Variations

Process Variations

Channel length
Oxide thickness
Dopant concentration
Threshold voltage

Temperature Variations

Mobility of electrons and holes
On-current

Variations

Can we isolate their effects?

How does this help?
Simulation Results for a Ring Oscillator

Temperature = 50°C  Process Corner = Low $V_t$  Nominal Delay = 151ps

Delay through SPICE simulations = 145.6ps
$\Delta$Delay = -5.4ps

Delay at $T=50^\circ$C & nominal process corner = 154.9ps
$\Delta$Delay = 3.9ps

Delay at low $V_t$ process corner & nominal temperature = 141.8 ps
$\Delta$Delay = -9.2ps

Delay using the principle of superposition = 145.7 ps
$\Delta$Delay = 3.9ps – 9.2ps = – 5.3ps
% Error = 1.8
Temperature Compensation

Determine the exact body bias voltage \( (v_{bn}, v_{bp})_{TABB} \) at each temperature point assuming ideal process conditions.

Performed using deterministic simulations

- Mathematically assisted TABB
- Bounded enumeration based TABB

Set of points which do not meet delay
Set of points which meet delay

- Delay increases
- Leakage decreases
Bounded Enumeration based TABB
Mathematically Assisted TABB

- Problem statement (NLPP):
  - Minimize $L(v_{bn}, v_{bp})$ s.t.
    - $D(v_{bn}, v_{bp}) \leq D^*$
    - $v_{bnmin} \leq v_{bn} \leq v_{bnmax}$
    - $v_{bpmin} \leq v_{bp} \leq v_{bpmax}$

- Need models for $L(v_{bn}, v_{bp})$ and $D(v_{bn}, v_{bp})$.

- Use 2nd order polynomial best fit expressions.

- Measure leakage and delay at sample points through deterministic simulations.

- Find the exact solution which lies along the blue line.

- Delay increases
- Leakage decreases

- Set of points which do not meet delay
- Set of points which meet delay

\[ v_{bn} \quad v_{bp} \]
\[ v_{bnmin} \quad v_{bpmin} \quad v_{bnmax} \quad v_{bpmax} \]
Solving the NLPP

\[ D(v_{bn}, v_{bp}) = D_0 \left[ \sum_{i=0}^{2} \left( \sum_{j=0}^{2} a_{ij} v_{bn}^j v_{bp}^i \right) \right] \]

Express delay as a product of two independent polynomials

\[ D = h(v_{bn}, v_{bp}) \approx f(v_{bn}) \cdot g(v_{bp}) \]
\[ f(v_{bn}) = 1 + x_1 v_{bn} + x_2 v_{bn}^2 \]
\[ g(v_{bp}) = 1 + y_1 v_{bp} + y_2 v_{bp}^2 \]

Eliminate one variable (say \( v_{bn} \))

Express \( L \) in terms of \( v_{bp} \) and find the minimum value using Newton Raphson method.
Comparing the 2 methods

- **Bounded Enumeration based TABB**
  - **Highlights**
    - Simple
    - Few computations
  - **Drawbacks**
    - Depends on the granularity of the voltages
    - Worst case complexity $O(n^2)$
    - Round-off error due to minimum voltage resolution

- **Mathematically assisted TABB**
  - **Highlights**
    - No elaborate search
    - Does not depend on the voltage resolution
    - Exact solution (which can be added with PABB)
  - **Drawbacks**
    - 1% modeling error
    - Overkill for P-well processes
    - Can be slower than TABB (if search space is limited)
Process Compensation

- **Problem Statement:**
  - Determine the body bias pair \((v_{bn}, v_{bp})\) for each WID variational region of each die at room temperature.

Measure the delay and leakage of each WID variational region at room temperature.

- \(D \geq D^*\) → **Apply Forward Body Bias**
- \(D \leq D^*\) → **Apply Reverse Body Bias**
Adaptive Body Bias (ABB)

Accounts for WID variation
PABB (Process Adaptive Body Bias)

- **v_{bn}**
- **v_{bp}**

Critical Path Replica

Measure delay & leakage

Build models for delay and leakage

Determine \((v_{bn}, v_{bp})_{PABB}\)

Formulate NLPP

\[
\begin{align*}
\text{Min } L(v_{bn}, v_{bp}) \text{ s.t. } \\
D(v_{bn}, v_{bp}) &\leq D^* \\
v_{bnmin} &\leq v_{bn} \leq v_{bnmax} \\
v_{bpmin} &\leq v_{bp} \leq v_{bpmax}
\end{align*}
\]

Performed at nominal (room) temperature
Simulation Set-up for PTABB

- **ISCAS85 benchmarks used.**
  - Simulations performed with BPTM 100nm technology and $V_{dd}=1.0V$.

- **Synthesis performed using SIS.**
  - Library of 26 gates (10 NOT gates, 5 NAND2 gates, 5 NOR2 gates, 3 NAND3 gates and 3 NOR3 gates) of different sizes.

- **Each benchmark placed in a different WID zone**
  - Can be independently compensated.

- **Bias range (-0.5 to 0.5V) for $v_{bn}$ and $v_{bp}$.**
  - $v_{step} = 0.1V$
# Results of TABB

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>T (C)</th>
<th>D* (ns)</th>
<th>No ABB</th>
<th>Enum TABB</th>
<th>Math TABB</th>
<th>Run time Ratio (t_{Math}/t_{Enum})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dly (ns)</td>
<td>v_{bn} (V)</td>
<td>v_{bp} (V)</td>
<td>v_{bn} (V)</td>
</tr>
<tr>
<td>C432</td>
<td>50</td>
<td>0.902</td>
<td>0.941</td>
<td>0.2</td>
<td>0.13</td>
<td>0.51</td>
</tr>
<tr>
<td>C432</td>
<td>75</td>
<td>0.986</td>
<td>11.2</td>
<td>0.4</td>
<td>0.36</td>
<td>1.63</td>
</tr>
<tr>
<td>C880</td>
<td>50</td>
<td>0.763</td>
<td>0.801</td>
<td>0.2</td>
<td>0.16</td>
<td>0.52</td>
</tr>
<tr>
<td>C880</td>
<td>75</td>
<td>0.838</td>
<td>6.85</td>
<td>0.5</td>
<td>0.49</td>
<td>3.11</td>
</tr>
<tr>
<td>C1355</td>
<td>50</td>
<td>0.83</td>
<td>0.841</td>
<td>0.2</td>
<td>0.17</td>
<td>0.55</td>
</tr>
<tr>
<td>C1355</td>
<td>75</td>
<td>0.879</td>
<td>11.9</td>
<td>0.5</td>
<td>0.5</td>
<td>3.10</td>
</tr>
<tr>
<td>C3540</td>
<td>50</td>
<td>1.33</td>
<td>1.39</td>
<td>0.2</td>
<td>0.19</td>
<td>0.41</td>
</tr>
<tr>
<td>C3540</td>
<td>75</td>
<td>1.45</td>
<td>37.5</td>
<td>0.3</td>
<td>0.37</td>
<td>0.89</td>
</tr>
<tr>
<td>C5315</td>
<td>50</td>
<td>1.20</td>
<td>1.25</td>
<td>0.2</td>
<td>0.19</td>
<td>0.42</td>
</tr>
<tr>
<td>C5315</td>
<td>75</td>
<td>1.30</td>
<td>35</td>
<td>0.3</td>
<td>0.40</td>
<td>1.17</td>
</tr>
<tr>
<td>C6288</td>
<td>50</td>
<td>3.64</td>
<td>3.82</td>
<td>0.2</td>
<td>0.17</td>
<td>0.47</td>
</tr>
<tr>
<td>C6288</td>
<td>75</td>
<td>3.99</td>
<td>57.7</td>
<td>0.4</td>
<td>0.37</td>
<td>1.75</td>
</tr>
</tbody>
</table>
Process Compensation

- PABB actually performed using post-silicon tuning.
  - Simulations provide an overview of the utility of our method.
- Test structure (critical path replica) is a RO simulated using 100nm BPTM.
  - RO used to determine Go- No Go for each WID.
- Monte Carlo simulations (50) done using Gaussian distributions for $v_{tn0}$ and $v_{tp0}$.

Simulations show that the yield is $\approx 50\%$ at room temperature and decreases gradually with increase in $T$. 
Process Compensation

- For each WID region, for each die, calculate the voltages \( (v_{bn}, v_{bp}) \) by solving the NLPP.
- For simulation purposes, we assume that all WID regions have the same \( v_{tn0} \) and \( v_{tp0} \) distribution.
- One set of simulations on the RO can be extended to all ISCAS benchmarks.

<table>
<thead>
<tr>
<th></th>
<th>NMOS-RBB</th>
<th>NMOS-RBB</th>
<th>NMOS-FBB</th>
<th>NMOS-FBB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PMOS-RBB</td>
<td>PMOS-FBB</td>
<td>PMOS-RBB</td>
<td>PMOS-FBB</td>
</tr>
<tr>
<td>No of dies</td>
<td>6</td>
<td>42</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
## PTABB Compensation

<table>
<thead>
<tr>
<th></th>
<th>Temp</th>
<th>%Yield</th>
<th>Accepted Dies (out of 50)</th>
<th>P-FBB N-RBB</th>
<th>P-RBB N-RBB</th>
<th>P-FBB N-FBB</th>
</tr>
</thead>
<tbody>
<tr>
<td>C432</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>35</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>C432</td>
<td>75</td>
<td>76</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>C880</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>24</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>C880</td>
<td>75</td>
<td>54</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>C1355</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>24</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>C1355</td>
<td>75</td>
<td>48</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>C3540</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>C3540</td>
<td>75</td>
<td>92</td>
<td>46</td>
<td>0</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>C5315</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>29</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>C5315</td>
<td>75</td>
<td>76</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>C6288</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>27</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>C6288</td>
<td>75</td>
<td>78</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>39</td>
</tr>
</tbody>
</table>
Conclusion

- Bidirectional Adaptive Body Bias can be used to improve the yield of dies for reasonable ranges of operating temperatures.
- New scheme to determine the exact values of body bias, **PTABB compensation** developed.
- One-time compensation for process variations and run-time compensation for temperature variations performed.
Backup
Independence of Variations

- Delay of a combinational circuit in the presence of temperature and process variations: \( D = f(x, T) \)
  - \( x \) is the vector of process variables
  - \( T \) is the operating temperature.
- \( x_0 \) and \( T_0 \) are the nominal values: \( f(x_0, T_0) = D^* \).
- At any other point \( (x_1, T_1) \) \( \Delta D \) can be written as
  \[ \Delta D = f(x_1, T_1) - f(x_0, T_0) \]
- If \( x \) and \( T \) are independent variables,
  \[
  \Delta D \approx \left[ f(x_1, T_0) - f(x_0, T_0) \right] + \left[ f(x_0, T_1) - f(x_0, T_0) \right]
  \]
L_{\text{min}} = \emptyset; \text{ Set initial solution to maximum bias}

Apply NMOS Bias

No

Apply maximum PMOS bias

No

Is delay OK?

No

Apply PMOS bias

No

Reduce NMOS bias

Is delay OK?

No

Is leakage OK?

No

Reduce PMOS bias

Update Solution

Return Final Soln

Final Solution
(v_{bn},v_{bp})_{PTABB} = (v_{bn},v_{bp})_{PABB} + (v_{bn},v_{bp})_{TABB}

### Temperature

<table>
<thead>
<tr>
<th>Temp</th>
<th>v_{bn}</th>
<th>v_{bp}</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PTABB values for each circuit block determined through simulations.

PABB values for each circuit block determined through post-silicon tuning.

PTABB values programmed into the ROM for each WID variational region.
# Results of Simulation on RO

Simulation using BPTM 100nm model files

RO delay with nominal temperature and process conditions = 151ps

<table>
<thead>
<tr>
<th>Temp</th>
<th>Process Corner</th>
<th>DelayPT $f(x_1, T_1)$</th>
<th>DelayP $f(x_1, T_0)$</th>
<th>DelayT $f(x_0, T_1)$</th>
<th>$\Delta$DelayPT</th>
<th>$\Delta$DelayP $+ \Delta$DelayT</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Low $V_t$</td>
<td>145.6</td>
<td>141.8</td>
<td>154.9</td>
<td>-5.4</td>
<td>-5.3</td>
<td>-0.1</td>
</tr>
<tr>
<td>50</td>
<td>Low $V_t$</td>
<td>165.3</td>
<td>161.2</td>
<td>154.9</td>
<td>14.3</td>
<td>14.2</td>
<td>0.1</td>
</tr>
<tr>
<td>75</td>
<td>High $V_t$</td>
<td>149.2</td>
<td>141.8</td>
<td>158.6</td>
<td>-1.8</td>
<td>-1.5</td>
<td>-0.3</td>
</tr>
<tr>
<td>75</td>
<td>High $V_t$</td>
<td>169.3</td>
<td>161.2</td>
<td>158.6</td>
<td>18.3</td>
<td>17.8</td>
<td>0.5</td>
</tr>
</tbody>
</table>