High-Performance SRAM Design

Rahul Rao

IBM Systems and Technology Group
Thought exercise

WWL

RWL

WBL

WBLb

RBL
Implement logic function via 8T merge

- Concept: Use the 8-T portion of the cell for implementing logic functions
  - Possible due to decoupling of read and write paths

- RBL discharges when either C0 or C1 is high
  - Read stack remains 2-high

- And function: switch definition of WBL and WBLB

- An OR2 embodiment is shown on right
  - Can be complex OR4, OR8, etc.
A More Practical Implementation

+ Share RWL between adjacent cells  
  => 3 word-lines in 4 metal tracks  
  – Reduces coupling capacitance (+ performance and power)
+ All FEOL features identical to conventional 8T cell  
  – OR  
  – upper device in stack can be made smaller, reducing cell size
The Phoenix Processor: A 30pW Platform for Sensor Applications

(S. Hanson, VLSI Symp '08)
Multi-porting SRAM Cell

WL0a

WL0b

BL0

BL1

BLB1

BLB0
Multi read ports
Multi read ports

RBL2
WWL
WBL
WBLb
RBL1
RWL0a
RWL0b
Figure 2.1 SRAM architecture

During the read operation the integrated SA on each column (sometimes shared between more columns) will be employed to read the data. In write operation, the write drivers will force the BL and BLB of selected column to '0' or '1' and the input data will be written into the internal nodes of the selected cell.

Hence, a typical column of SRAM consists of the following blocks:

- **Bank and Bank Conflicts**
- **Row Decoder**
- **Column Decoder**
- **Sense Amplifier & Write Driver**
- **Global Data Bus**
- **Global Read/Write**
- **Address Buffer**
- **Timing & Control**
- **Precharge Circuit**
- **2^n bits**
- **2^n x 2^n cell**
- **CELL**
- **BL, BLB**
- **Address**
- **R/W, CS**
- **Blocks**

Signals is used for the determination of read or write operation and the chip set (CS) signal is usually employed in multi-chip designs.
Global and Local Variations

Random Dopant Fluctuation

Global

Intra-die

Inter-die

Local
Hold Failure

WL = 0
L = '1'
R = '0'

V_DDH

Time >

Voltage

WL
V_DDH
V_R
V_L

Time ->

Voltage

WL
V_DDH
V_R
V_L

S. Mukhopadhyay, ITC 2010
Read Failure

Voltage

Time -->

S. Mukhopadhyay, ITC 2010
Write Failure

Voltage vs. Time:
- WL
- VR
- VL

Diagram:
- AXL
- L='1'
- PL
- PR
- NL
- NR
- R='0'
- BL
- BR

Time ->
Inter-die Variation & Cell Failures

Inter-die Vt shift ($\Delta V_{\text{th,GLOBAL}}$)

Low-Vt Corners
- Read failure $\uparrow$
- Hold failure $\uparrow$

High-Vt Corners
- Access failure $\uparrow$
- Write failure $\uparrow$

Failures in SRAM Array

Overall Cell Failure:

\[ P_F = P[\text{Fail}] = P[A_F \cup R_F \cup W_F \cup H_F] \]

- **\( P_{COL} \):** Probability that any of the cells in a column fail

\[ P_{COL} = 1 - (1 - P_F)^{N_{ROW}} \]
Transistor Sizing

- Slide contributed by K. Roy, Purdue
Larger redundancy
(1) more column to replace (less memory failure).
(2) smaller cell area (larger cell failure).
Question

- **Array redundancy**
  a) Improves cell stability
  b) Degrades cell performance (i.e., increases read and write times)
  c) Does not require any change to cell peripheral circuits
  d) Row redundancy is better than column redundancy
Example: Multi-VCC for SRAM Cell

- Create differential voltage between WL and Cell to decouple the Read & Write
  - Write: $V_{WL} > V_{Cell}$
  - Read: $V_{WL} < V_{Cell}$

Source: K. Zhang et al. ISSCC 2005
Dynamic Circuit Techniques for Variation Tolerant SRAM

\[ V_{WL} = V_{DD} + \Delta \]

\[ V_{cell} = V_{DD} - \Delta \]

\[ V_{BL} = 0 - \Delta \]

\[ V_{BR} = V_{DD} \]

<table>
<thead>
<tr>
<th></th>
<th>Read</th>
<th>Write</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{WL} )</td>
<td>Lower ( V_{WL} ) =&gt; lower ( V_{read} ) (weak AX)</td>
<td>Higher ( V_{WL} ) =&gt; Strong AX helps discharge</td>
</tr>
<tr>
<td>( V_{cs} )</td>
<td>Higher ( V_{cs} ) =&gt; lower ( V_{read} ) (strong PD) Higher ( V_{trip} )</td>
<td>Lower ( V_{cs} ) =&gt; Weak PUP</td>
</tr>
<tr>
<td>( V_{BL} )</td>
<td>Weak impact</td>
<td>Negative ( V_{BL} ) for 0 =&gt; strong AX helps discharge</td>
</tr>
</tbody>
</table>
Example: Dual-Vcc based Dynamic Circuit Techniques

- Dynamic VCC MUX is integrated into subarray
- VCC selection is along column direction to decouple the Read & Write

Source: K. Zhang et. al. ISSCC 2005
Negative Bit Line Scheme

Conventional

This Scheme

Δ ≈ C_{boost}/C_{BL}

BIT_EN

NSEL

CS

WR

PCHG

BIT_EN generating block

NSEL

V_{in} C_{boost}

V_{BL}

C_{BL}

DB=“1”

D=“0”

Source: S. Mukhopadhyay, R. Rao et. al, TVLSI 2009
Effectiveness Considerations: Writability improvement

- Various dynamic schemes have different effectiveness in improving writability for similar read stability
  - Higher $V_{WL}$ is most effective

Source: S. Mukhopadhyay, R. Rao et. al, TVLSI 2009
Impact on Active Data-Retention

- Column based read-write control adversely impact the active data-retention failures
  - DC negative bitline has higher active data-retention failures
  - Tran-NBL and lower $V_{cs}$ have comparable failure rates

Source: S. Mukhopadhyay, R. Rao et. al, TVLSI 2009
Dynamic Circuit Techniques for Variation Tolerant SRAM

\[ V_{WL} = V_{DD} + \Delta \]
\[ V_{cell} = V_{DD} - \Delta \]

| \( V_{BL} = 0 \) - \( \Delta \) | \( V_{BL} \) | \( V_{BR} = V_{DD} \) |

<table>
<thead>
<tr>
<th>Read</th>
<th>Write</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower ( V_{WL} ) =&gt; lower ( V_{read} ) (weak AX)</td>
<td>Higher ( V_{WL} ) =&gt; Strong AX helps discharge</td>
</tr>
<tr>
<td>Higher ( V_{cs} ) =&gt; lower ( V_{read} ) (strong PD) Higher ( V_{trip} )</td>
<td>Lower ( V_{cs} ) =&gt; Weak PUP</td>
</tr>
<tr>
<td>Weak impact</td>
<td>Negative ( V_{BL} ) for 0 =&gt; strong AX helps discharge</td>
</tr>
</tbody>
</table>
Implementation Consideration: Half-Select Stability

\[ V_{\text{cell}} = V_{\text{DD}} - \Delta \]

- Higher \( V_{\text{WL}} \)
  - Row-based scheme
  - Degrades half-select read stability of the unselected columns

- Lower \( V_{\text{cell}} \) or negative bit-line
  + Column-based scheme
  + Half-select read stability remains same
Assist Methods

**WLOD** *(WL Overdrive)*
- Strengthen PG

**VCDL** *(V\text{DD,CELL} Lowering)*
- Weaken PU

**NBL** *(Negative BL)*
- Strengthen PG

**SBL** *(Suppressed BL)*
- Weaken BL noise

**WLUD** *(WL Underdrive)*
- Weaken PG
Proposed Dual-Transient WL (DTWL)

- Dual-Transient WL (DTWL) controls WL transiently
- DTWL provides mix-up assist for read and write
  - Covers different process-corner
  - DTWL mitigate the impact of WLOD (WA)
Question

- Of the various assist methods
  a) Negative bit line scheme does not help 8-T sram cell
  b) Word line under drive does not help 8-T sram cell
  c) Word line over drive does not help 7-T conditionally decoupled sram cell
  d) VCDL does not help any kind of asymmetric sram cell
Precharge Time (Timing diagrams)

SRAM as random number generator