Announcements

- Homework #4 due today
- Final exam on May 8 in class
- Project presentations on May 3, 1-5pm
Class Material

- Last lecture
  - Technology and environment variability
- Today’s lecture
  - SRAM

SRAM Scaling Trends

- Individual SRAM cell area able to track ITRS guideline
- Array area deviates from ITRS guideline at 90nm
- Memory design no longer sits on the 0.5x area scaling trend!
Memory Scaling

- Memory latency demands larger last level cache (LLC)
- Memory is more energy-efficient than logic
- LLC approaches 50% chip area for desktop and mobile processors
- LLC approaches 80% chip area for server processors

Vivek De, Intel 2006

6-T SRAM Cell

- Improve CD control by unidirectional poly
- Relax critical layer patterning requirements
- Optimizing design rules is key

Vivek De, Intel 2006
SRAM cell design trends

- Improve CD control by unidirectional poly
- Relax critical layer patterning requirements
- Optimizing design rules is key
- Shorter bitline enables better cycle time and/or array efficiency
- Full metal wordline with wider pitch achieves better RC

Ion/Ioff: Cell Read and Leakage

[Graph showing cell read current and leakage current as a function of technology node (180nm to 65nm)].

H. Pilo, IEDM 2006
SRAM Cell/Array

- Read stability
- Write stability
- Read current

Access Transistor

Pull down
Pull up

SRAM Design – Hold (Retention) Stability

- Scaling trend:
  - Increased gate leakage + degraded $I_{ON}/I_{OFF}$ ratio
  - Lower $V_{DD}$ during standby
  - PMOS load devices must compensate for leakage
The Data-Retention Voltage (DRV) of SRAM

When \( V_{dd} \) scales down to DRV, the Voltage Transfer Curves (VTC) of the internal inverters degrade to such a level that Static Noise Margin (SNM) of the SRAM cell reduces to zero.

Qin, ISQED’04

DRV Condition:
\[
\left. \frac{\partial V_1}{\partial V_2} \right|_{\text{cell inverter}} = \left. \frac{\partial V_1}{\partial V_2} \right|_{\text{right inverter}}, \text{ when } V_{dd} = \text{DRV}
\]

Monte-Carlo Simulation of DRV Distribution

Histogram of cell #

Simulated DRV of 1500 SRAM cells (mV)
Read Stability – Static Noise Margin (SNM)

- Read SNM is typically the most stringent constraint
- SNM shrinks with each generation

Read Stability – N-Curve

- A, B, and C correspond to the two stable points A and C and the metastable point B of the SNM curve
- When points A and B coincide, the cell is at the edge of stability and a destructive read can easily occur

[Bhavnagarwala, IEDM’05]
Write Stability – Write Noise Margin (WNM)

- Write stability is becoming more stringent with scaling
- Optimizing read and write stability at the same time is difficult

[1] A. Bhavnagarwala, IEDM 2005
Write Stability – Traditional Write Margin (TWM)

- Highest BL voltage under which write is possible when BLC is kept precharged (left)
- Difference between VDD and lowest WL voltage under which write is possible when both bit-lines are kept precharged (right)
- Can be directly measured in large memory arrays via BL currents

Write Stability – Write Current (N-Curve)

- When biased as shown, the minimum current into a node provides a measure for the write-ability of the cell [1]
  - Indicates relative PG & PU drive strength
  - Must be positive to write a zero to that half of the cell
  - "Write-ability current" $I_{pr}$ = minimum current from both halves of the cell

The Conflict Between Read and Write

SRAM Design – Read/Write Stability

- Read margin is typically the most stringent constraint
- Cell read voltage must stay below cell trip voltage
  - Harder to achieve with process induced variations
  - Noise margin degraded with technology scaling
Bit-Line Load Impact on Stability

- Short Bit-Lines
  Bitline load: 32-64 cells per bitline

- Long Bit-Lines
  Bitline load: 128-512 cells per bitline
Multi-Voltage SRAM

<table>
<thead>
<tr>
<th></th>
<th>Read</th>
<th>Write</th>
<th>Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periphery</td>
<td>Vmin</td>
<td>Vmin</td>
<td>Vmin</td>
</tr>
<tr>
<td>BL Precharge</td>
<td>Vmin</td>
<td>Vmin</td>
<td>Vmin</td>
</tr>
<tr>
<td>WL</td>
<td>Vmin</td>
<td>Vmax</td>
<td>N/A</td>
</tr>
<tr>
<td>Cell V_{DD}</td>
<td>Vmax</td>
<td>Vmin</td>
<td>Vmin</td>
</tr>
<tr>
<td>Cell well</td>
<td>Vmax</td>
<td>Vmax</td>
<td>Vmax</td>
</tr>
</tbody>
</table>
Dynamic $V_{DD}$ Implementation

- VCC selection is along column direction to decouple the Read & Write

Zhang, ISSCC'05

The Rest in Class Presentations

- Read/Write assist circuits
- Alternate cells for subthreshold operation
- FinFET/double-gate designs
- Column design techniques
- Leakage suppression
- Sense amps...
SRAM Scaling

- Approaching fundamental limits:
  - Don’t scale cell size
  - Increase transistor count (from 6)
  - Change technology (e.g. double-gate FETs)
  - eDRAM
  - Or something else...

Other SRAM Alternatives

8-T SRAM [1]

- Dual-port read/write capability (register file like cells)
- N0, N1 separates read and write
  - No Read SNM constraint
  - Half-selected cells still undergo read stress – no single cell write capability
- Stacked transistors reduce leakage

eDRAM

- Process cost: Added trench capacitor

Next Lecture

- Latches and flip-flops