EE241 - Spring 2002
Advanced Digital Integrated Circuits

Lecture 27
Memory

References

- Rabaey, “Digital Integrated Circuits”
  - Gillingham, Evolution of DRAM
  - Nuhn, Cao, MacGillivray, DRAM Development Trends
- Chapter 14, Register Files and Caches, by R. Preston
- Chapter 15, Embedded DRAM, by Yamauchi and Yamada
## Semiconductor Memory Classification

<table>
<thead>
<tr>
<th></th>
<th>RWM</th>
<th>NVRWM</th>
<th>ROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Access</td>
<td>Non-Random Access</td>
<td>EPROM</td>
<td>Mask-Programmed Programmable (PROM)</td>
</tr>
<tr>
<td>SRAM</td>
<td>FIFO</td>
<td>E²PROM</td>
<td></td>
</tr>
<tr>
<td>DRAM</td>
<td>LIFO</td>
<td>FLASH</td>
<td></td>
</tr>
<tr>
<td>Shift Register</td>
<td>CAM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Array-Structured Memory Architecture

Problem: ASPECT RATIO or HEIGHT >> WIDTH

- Amplify swing to rail-to-rail amplitude
- Selects appropriate word

<table>
<thead>
<tr>
<th>Row Decoder</th>
<th>Column Decoder</th>
<th>Sense Amplifiers / Drivers</th>
<th>Storage Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{K}$</td>
<td>$A_{K-1}$</td>
<td>$A_{K-1}$</td>
<td></td>
</tr>
<tr>
<td>$A_{K-1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{K-2}$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Input-Output (M bits)
Memory Cells

(a) DRAM  
(b) SRAM  
(c) EPROM  
(d) Mask ROM

Read-Write Memories (RAM)

• STATIC (SRAM)
  Data stored as long as supply is applied
  Large (6 transistors/cell)
  Fast
  Differential

• DYNAMIC (DRAM)
  Periodic refresh required
  Small (1-3 transistors/cell)
  Slower
  Single Ended
3-Transistor DRAM Cell

No constraints on device ratios
Reads are non-destructive
Value stored at node X when writing a “1” = V_{WWL}-V_{Tn}

1-Transistor DRAM Cell

Write: C_s is charged or discharged by asserting WL and BL.
Read: Charge redistribution takes place between bit line and storage capacitance

\[ \Delta V = V_{BL} - V_{PRE} = (V_{BIT} - V_{PRE}) \frac{C_s}{C_s + C_BL} \]

Voltage swing is small; typically around 250 mV.
DRAM

1T DRAM requires a sense amplifier for each bit line, due to charge redistribution read-out.

DRAM memory cells are single ended in contrast to SRAM cells.

The read-out of the 1T DRAM cell is destructive; read and refresh operations are necessary for correct operation.

Unlike 3T cell, 1T cell requires presence of an extra capacitance that must be explicitly included in the design.

When writing a “1” into a DRAM cell, a threshold voltage is lost. This charge loss can be circumvented by bootstrapping the word lines to a higher value than $V_{DD}$.

1-T DRAM Cell

(a) Cross-section

(b) Layout

Used Polysilicon-Diffusion Capacitance
Expensive in Area
Advanced 1T DRAM Cells

Cell Plate Si
Capacitor Insulator
Storage Node Poly
2nd Field Oxide
Trench Cell

Insulating Layer
Transfer gate
Isolation
Storage electrode
Refilling Poly
Si Substrate
Stacked-capacitor Cell

DRAM Evolution

  » Gillingham, Evolution of DRAM
  » Nuhn, Cao, MacGillivray, DRAM Development Trends
1K DRAM

- Intel 1103, late 1971.
- Cost effective - <1c/bit
- PMOS, silicon gate, 1M1P
- Vdd=0V, Vss=16V, Vbb=20V
- 300ns access, 580ns cycle, 2ms retention

4k DRAM

- Texas Instruments TMS4030, introduced 1973
- NMOS 1M1P, TTL I/O
- 1T Cell, Differential Sense Amp
- Vdd=12V, Vcc=5V, Vbb=-5V, Vss=0V
16k DRAM

- MOSTEK MK4116, introduced 1977
- Address multiplex
- NMOS 2P1M
- \( V_{dd}=12\text{V}, V_{cc}=5\text{V}, V_{bb}=-5\text{V}, V_{ss}=0\text{V} \)

64k DRAM

- 1980, NMOS 2P1M
- 5V only, internal \( V_{bb} \) generator
- Boosted wordline, eliminate \( V_t \) loss
256K DRAM

- Folded bitline architecture
- NMOS 2P1M, poly 2 (polycide), gate, W/L, 5V
- Redundancy
- Active restore

1M DRAM

- CMOS, N-well, Vbb substrate, 3P1M
- Boosted circuits
- Vdd/2 bitline reference, Vdd/2 cell plate
4M DRAM

- CMOS 4P1M
- x16 output, self refresh
- 3D cell structures, stacked or trench

16M DRAM

- CMOS 4p2M, Vdd=5V, internally 3-4V
- EDO – Extended Data Out
Row Decoders

Collection of $2^m$ complex logic gates
Organized in regular and dense fashion

(N)AND Decoder

$$WL_0 = A_0 A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9$$

$$WL_{511} = \overline{A_0 A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8 A_9}$$

NOR Decoder

$$WL_0 = \overline{A_0 + A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 + A_9}$$

$$WL_{511} = \overline{A_0 + A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 + A_9}$$

A NAND decoder using 2-input pre-decoders

Splitting decoder into two or more logic layers produces a faster and cheaper implementation
Decoders

- Word line selection results in only one critical transition
- Skew the gates sizing
- Requires resetting for the next transition
- Precharging
- Self-resetting; Delayed-resetting

Dynamic Decoders

- Propagation delay is primary concern

Dynamic 2-to-4 NOR decoder 2-to-4 MOS dynamic NAND Decoder
Precharged Decoders

- Most designs use NAND-based design
- NOR based design has a speed advantage, but big power problem
  » All outputs are initially precharged
  » When decoder evaluates, $2^{M-1}$ outputs discharge, only one stays high – power problem
- In NAND-based designs only one output changes
- Clock power is still very large: precharge and evaluate transistors switch every cycle

Self-Resetting Decoders

Chappell, JSSC 11/91
Self-Resetting Decoders

Park, ISSCC'98
4Mb DDR SRAM

Source-Coupled Logic

Static
Dynamic

Source-Coupled

Nambu, JSSC 11/98
Source-Coupled Logic

Conventional Decoder

Source-Coupled

Sense Amplifiers

Idea: Use Sense Amplifier
**Differential Sensing - SRAM**

(a) SRAM sensing scheme.

(b) Doubled-ended Current Mirror Amplifier

(c) Cross-Coupled Amplifier

**Latch-Based Sense Amplifier**

Initialized in its meta-stable point with EQ
Once adequate voltage gap created, sense amp enabled with SE
Positive feedback quickly forces output to a stable operating point.
Open Bitline Architecture

DRAM Read Process with Dummy Cell

(a) reading a zero

(b) reading a one

(c) control signals
Open Bit-line Architecture — Cross Coupling

Folded-Bitline Architecture
Transposed-Bitline Architecture

(a) Straightforward bitline routing.

(b) Transposed bitline architecture.

Alpha Particles

1 particle ~ 1 million carriers
Redundancy

Redundant rows

Redundant columns

Memory Array

Row Address

Fuse Bank

Column Decoder

Column Address

Redundancy and Error Correction

Percent yield

0 20 40 60 80 100

Average number of failing cells per chip

0 500 1000 1500 2000 2500 3000 3500 4000

ECC only

Redundancy only

Redundancy and ECC