Organization of a RAM.
Organization of a RAM.
Memory Array Architecture

- Each row address accesses both sides of the array
- Good for large capacity array not too

Full plane

- Each row address accesses one side of the array
- Lower power consumption

Half plane
• Block-oriented

* Reduce bit line length

Each block contains its own block selection (to select specific block) CRT, x-decoder and select one of the rows
sense amplifier.

• SRAM cell design

<table>
<thead>
<tr>
<th>BL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>φ</td>
<td></td>
</tr>
<tr>
<td>premage clock</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td></td>
</tr>
<tr>
<td>M6</td>
<td></td>
</tr>
<tr>
<td>BL</td>
<td></td>
</tr>
</tbody>
</table>

Bit-line loads

wont line WL

other cells
- Transistor sizing is very important.

- Example for Read

![Diagram]

- Assume the cell stores logic 1
  - $Q = 1$, $\overline{Q} = 0$

- Before read, both bit lines are precharged to 1.

- We expect $BL = 1$, but $\overline{BL}$ will be pulled down.

- Capacitance of $\overline{BL}$ and $BL$ are much larger (pF) than transistor capacitances, so the swing of $\overline{BL}$ will not be too large.

- sense the difference needs a sensor amplifier to amplify it!! the signal.
\[ V_{\overline{a}} = \left( \frac{V_{dd}}{R_{M5} + R_{M4}} \right) R_{M1} \approx 0.3 V_{dd} \]

- A reasonable sizing is

\[ W_5 = W_{\text{min}}, \quad W_1 = 2.5 W_{\text{min}}. \]

Similarly, \( W_6 = W_{\text{min}}, \quad W_3 = 2.5 W_{\text{min}}. \)

- Example for write

- If write 1, no change!

- So, let us try to write 0.
- BL cannot charge the state of $\bar{Q}$ as we discussed.

- $Q=1$ must be charged first by discharging through $M_6 \rightarrow BL$ line.

- $Q=1$ must drop below $U_{turpo}$ such that $M_2$ can be on. So, there is a sizing problem between $M_4$ and $M_6$.

- Once $Q$ drops to below $U_{turpo}$, $M_2$ begins to be on and charge $\bar{Q}$ to 1 and charge the state of the cell. (Get Stabilized)!!

- $M_4$ must be large enough, such that $Q$ can drop to small (in size) $i.e., i_t$ R must be large!!

\[
\left(\frac{W}{L}\right)_4 \leq 1.8 \left(\frac{W}{L}\right)_6
\]

\[\Rightarrow W_4 = W_2 = 1.8 \text{ Wmin}\]
Bit line Load.

- In fact, the bit lines can be charged to lower than \( V_{dd} \) to save power.

- However, if BL (in PMOS) is too low, \( Q=1 \) cannot be maintained, and the state may be changed.

\[ \text{Bit line voltage} \uparrow \Rightarrow \text{more stable memory state} \]

- can have many different designs.
No matter what kind of pull-up act, equalizing device must be added.

After each R/W operation, this must be activated for a while to equalize the voltage.

Row Decoder: $8 \rightarrow 256$ mapping (Example) - NAND-type

Evaluation

Precharge
Dynamic ck

\( \phi = 0 \), all outputs are zero.

\( \phi = 1 \), only one word line would be 1.

- Very power-saving, slower (\( \phi \) larger height of evaluation
  (only one line discharge)
- Must add inverter buffer, since there is no
data source to drive the high capacitive load.

\( \text{NOR-type} \)

3 \( \rightarrow \) 8 mappig

\[ A_1 = A_2 = A_3 = 0 \]
\[ \Rightarrow z = 1, \]
\[ \Rightarrow y = 0, \]
\[ x = 1 \]
- Faster, but consumes more power
  - N-1 rows are disabling.

Evaluation

- Multi-stage decoder (or called Predecker)
  - Multi-stage decoder (Compromise between NOR-type & NAND-type)

NAND implementation details in p. 148

Group
2^{m-1}

- Only one stage be activated
  - Save power!!!

NAND implementation

- Can be partitioned to more stages
Decoder: Show by 2→4 mapping

Don't need to be fast!

This is near the end of read operation.

Tree decode

Slow!
May need to pass many transistors.
Sense Amplifier

1. Sense Amplifier

2. Voltage equalizer

3. Sram cells

4. These two invert push each other!!!
• Inverter is a high-gain device

- Charge in input 0.5V \rightarrow output charge about 2.0V
- \sim 4 \text{ times of gain.}

• Initially, both BL and BL are precharged to \sim 1.5V.

• If BL is discharged to 1.0V, \overline{BL} \text{ remains at 1.5V} \rightarrow 2.5V
  \overline{BL} \text{ remains at 1.5V} \rightarrow \overline{BL} \text{ dropped to 0.5V} \rightarrow 0V.

• If \overline{BL} is discharged to 1.0V, \overline{BL} \text{ remains at 1.5V} \rightarrow 2.5V
  \overline{BL} \text{ remains at 1.5V} \rightarrow \overline{BL} \text{ dropped to 0.5V} \rightarrow 0V, \text{ eventually}

• This is most commonly used in real designs and
  consume much less power than current mirror based S.A.

• The sense amplifier enable signal is generated using a
  self-timed approach to reduce the power consumption.
  That is BS will not charge to 2.5V \rightarrow \overline{BL} \text{ will not be discharged to 0V} \rightarrow \overline{BS} \text{ or vice versa.
Read/write operation
Low-power memory design

1. Memory Sub-banking
   - R rows, C columns, Cell: C (bit line switch/Cell)
   - C totally, R.C. Cell of cap. is switched.
   - Try to partition the memory

![Diagram of memory sub-banking]

Make it B banks

- if interconnection is not considered.
- can find an optimal solution.
Divided word line

- Needs at least 2 metal lines
- Local word line: polysilicon
- Global word lines: metal 1
- Bit lines: metal 2.
- Divided Word Line Architecture

- Reduce power consumption in Bit Lines.
• Bit line segmentation

• Reduce voltage swing on Bit Line

* Not quite feasible

... affect noise sensitivity, complexity of sense amplifier, performance of the RAM core.

[Diagram of bit line and word line waveforms]

Complete cycle (without word line termination)
- Pulsed word lines

- Try to limit the bit line voltage discharge.
- Try to enable the word lines for precisely the time needed to develop the bit cell voltage discharge.

- Odd # of inversions

- It is hard to control the width
- Try to use self-timed word lines.
- Self-timed word lines.

- Fast column 'close to word line drivers.'

- Sense amplifiers

1. SR flip-flop is set, and the **word line** is triggered. (SA)

2. When the dummy SA (slowest) generate 1, the rest of the columns would have been sensed.

3. The high signal reset the flip-flop and turn off the **word line**. (SA)
Summary: power optimization for SRAM

Reduce total capacitance switched

1. Banked organization
2. Word-line division
3. Multi-stage decoder
4. Bit line segmentation

Reduce the voltage swing across switched capacitances

1. Self-timed RSRAM
   - turn off word line & SA enable, s.t. the voltage swing in the bit lines can be minimized.
2. Reduce bit line voltage swing.