Lecture 18

SRAM Cell and Column I/O Design

Res Saleh
Dept. of ECE
University of British Columbia
res@ece.ubc.ca

Architecture of 64Kb SRAM
SRAM Storage Cell

Uses six transistors (called 6T memory cell):

- Read and write operations use the same port. There is one wordline and two bit lines. The bit lines carry complementary data, a fact that will be used to reduce access time. The cell layout is small since it has a small number of wires (but large relative to DRAM).

CMOS SRAM Cell Design

Cell Design

Problem: Find wa, wd, wp such that
1) minimize cell area
2) obtain good read and write cell margins
3) good soft error immunity
4) good cell read current

in that order

Since the cell is symmetric we need only design three transistor sizes
CMOS SRAM cell design - read

- Make sure that internal node $a$ does not go high enough to turn on $M_2$
- Use threshold voltage as a maximum allowable voltage at internal node during read; Make current ratio between $M_1$ and $M_3$ about 3 to 4
- Want to design device sizes such that read current is high enough to create desired differential voltage on bit lines $\sim 200mV$ within a specified amount of time

Rule-of-Thumb: $W_1/W_3 = 1-1.5$

CMOS SRAM cell design - write

- Need to make sure $M_4$ is strong enough to pull internal node $a$ low while $M_6$ is trying to pull it high
- Use switching threshold as trigger point for regenerative switch point; usually want to make it less than the switching threshold
- This will force inverter $M_5$-$M_1$ switch to new state
- Make ratio of currents between $M_4$ and $M_6$ about 3 to 4

Rule-of-Thumb: $W_4/W_6 = 1-1.5$
SRAM Cell Layout

- Word line running horizontally
- Bit lines running vertically
- Cross-coupled inverters on top
- Access transistors on bottom

Portion of the core array using SRAM cell
- Scaled in dimensions compared to single bit cell
- 2 cells across, 3 cells high
- Replicated in this manner to build the core array

Wordline Capacitance

- Word line presents a large capacitance to the decoder
- Each cell loads the word line with two transistor capacitances and one wire capacitance (plus wire resistance)
- Total capacitance is the capacitance per cell x number of cells on word line
Column I/O Operation

Circuits that perform read and write on the array are column I/O

- Bitline load
  - Can be static or precharged
  - Proper configuration depends on amplifier design
- For read
  - Bit lines must start at around Vdd
  - Swings should be small for fast operation
  - Involves Mux and sense amplifier design
- For write
  - Need to drive one of the bitlines to Gnd
  - Mux and write driver design
  - Often use different I/O lines for read and write

Bitline Capacitance

Load capacitance is mostly self-loading of the cells

- Drain cap and drain contacts (0.5-1 fF) of transistors are shared
  - Junctions are biased at Vdd (lower cap than normal)
- Wire capacitance ~ .2fF/micron of wire

![Diagram of memory array with bitlines and wordlines](image)
Bitline Load Options

Important to equalize bitline voltage before reads

Latch-based Sense amplifier

Analog differential Sense amplifier

Write Circuitry

- Precharge bitlines high
- Pull one column line low
- Turn on word line
- Wait until internal values of cell are established
- Turn off word line
- Design precharge transistors to pull bit lines high
- Design write drivers to pull one side low depending on data value
Read Circuitry

- Precharge bitlines high
- Turn on word line
- One line will slowly discharge
- Wait until bit line reaches required low voltage level
- Turn on column select
- Amplify difference with sense amplifier
- Design sense amplifier based on desired response time and power requirements
- Use precharge based on type of sense amp used.

Column Decoder/Mux

- Need decoder for column address followed by a mux to select column for input or output operations
- Require two outputs to drive complementary pass gates
- Since the requirements for read and write are different, can use separate read and write IO lines
- Have PMOS access for the read IO lines, since the read happens near Vdd
- Have NMOS devices for the write IO lines, since you need to drive bitlines to Gnd (see next slide)
Column Muxing – Separate I/O

Multi-Level Column Decoding

• Alternatives for column selection are tree decoder, regular decoder + pass transistor, or some combination of the two
  
  • Shown on the right is a tree decoder
    - switches driven directly by address bits and their complements
    - total of \(2^{M-1}\) devices
    - large devices to reduce resistance
    - long paths \(\rightarrow\) large \(C\)
      \(\rightarrow\) SLOW

  • To speed up, add buffers or use adjust sizing of devices
Building Amplifiers

- We need an amplifier to handle small voltage swings on the bit lines for fast operation
- Normally you need to choose between
  - Drawing DC power (diff. sense amplifier)
  - Using a clock edge (latch-based amplifier)
    (to turn DC power on only when the signal is present)
- For CMOS logic gates
  - When input is at VDD or Gnd, one of the transistors is off
  - Nothing can happen until this transistor turns on
    • And even then you need to wait some more for gate to switch
    • Sitting in low gain region of transfer curve
- For an amplifier
  - Want to be in high-gain region (saturation)
Latch-based Sense Amplifier

- It must sense a very small signal
- It must consume a small area
  - Need one for each bitline
  - Or sets of bitline (4 or 8)
- Simplest design:
  - Two back-to-back inverters
  - Add a clocked pulldown
- Once Bit and Bit_b are established, turn on pulldown device to activate inverters
- Side with lower voltage will drop to 0V while the other side stays high

Using Clocks and Regeneration

- Three stages of operation
  - Precharge
  - Sample
  - Regenerate
- At the end of sample
  - Small bitline voltage on sense and sense_b
- Regenerate
  - M2 and M3 turn on
  - Voltage difference causes current difference
  - Which causes larger voltage difference