**CHAPTER 3**

**DESIGN CONCEPT AND APPROACH**

In this chapter the concepts of 6T and 4T SRAM cells are discussed. Different operations (read, write and hold) taking place, design considerations and the working in detail is explained for both the SRAM cells.

**3.1 WORKING**

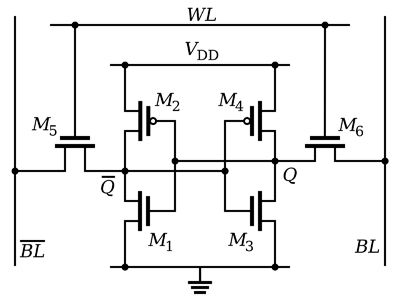
SRAM cell works in three operating modes.

* **Hold state:** In this state data is stored in the cell for future use.
* **Read state:** In this state, data stored in the cell is brought onto the bit-lines
* **Write state:** In this state, data on the bit-lines is written into the cell

The working of 6T and 4T SRAM cells in detail is discussed as follows.

**3.1.1 WORKING OF 6T SRAM CELL**

6T SRAM cell uses a simple bi-stable latch circuit to hold a data bit. A pair of cross coupled inverters provides the storage while two access transistors (NFETs) provide read and write operations. The access transistors are controlled by the word line signal WL that defines the operational modes.



**Fig (3.1) 6T SRAM Cell**

The cell works in three operating modes.

**Hold state:** When WL=0, both access FETs are OFF and the cell is isolated. This defines the “hold’’ condition and the cell retains the value stored. Both bit and bit-bar lines are pre-charged to VDD during this state.

**Write operation:**

To perform ‘‘write’’ operation, the word line is brought up to a value of WL=1. This turns ON the access transistors and bit, bit-bar lines get connected to the write circuitry. Value to be written is placed on bit line and its complimentary value is placed on bit-bar line. Pre-charge remains off during write operation.

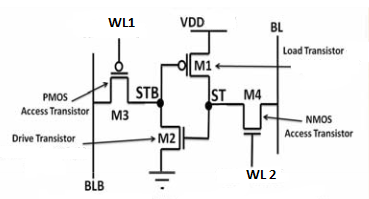
When logic ‘1’ is written into the cell, bit line is made high and bitbar line is made low. Referring to the Fig (3.1), access transistors M5 and M6 are turned ON. Node Q is charged to VDD which turns M1 ON and M2 OFF. M1 pulls Q**—** node to ground. M3 turns OFF and M4 is turned ON which pulls node Q to VDD.

When logic ‘0’ is written into the cell, bit line is made low and bitbar line is made high. Referring to the Fig (3.1), access transistors M5 and M6 are turned ON. Node Q is pulled down to logic ‘0’ value which turns M2 ON and M1 OFF. M2 pulls Q**—** node to VDD. M4 turns OFF and M3 is turned ON which pulls node Q to ground.

**Read operation:** During the read operation word-line is made high, WL=1. This turns on the access transistors and bit, bit-bar lines get connected to read circuitry. Pre-charge remains OFF during read operation. Bit and bitbar lines act as outputs and are fed into a sense amplifier that determines the stored state.

**3.1.2 WORKING OF 4T SRAM CELL**

4T SRAM cell consists of load and driver transistors and two access transistors. Unlike in 6T cell where both the access transistors are NFETs, in 4T cell one of the access transistors is a PFET. Two different word lines WL1 and WL2 are used for controlling PFET and NFET access transistors respectively. Circuit diagram of 4T SRAM cell is shown in the Fig (3.2).

.

**Fig (3.2) 4T SRAM Cell**

The cell works in three operating modes.

* **Hold state:** When WL1= 0V and WL2= 0.3V, both access transistors are OFF and the cell is isolated. This defines the “hold’’ condition and the cell retains the value stored. Different pre-charge voltages are used for bit and bit-bar lines. Bit line is pre-charged to VDD and bit-bar line is pre-charged to ground during this state.
* **Write operation:**

When a write operation is issued the memory cell will go through the following steps.

**1) Bit-line driving:** For a write, data is placed on BL, and then WL2 asserted to VDD, but voltages on word-line1 and BLB maintained at idle mode ( WL1= 0.25V and VBLB=VDD). Only pre-charge on bit-line is turned OFF.

**2) Cell flipping:** This step includes two states as follows.

**(a) When data is logic ‘0’:** In this state, ST node is pulled down to GND by NMOS access transistor, and therefore the drive transistor will be OFF. STB node will be floated and then pulled up to voltage of BLB (VDD) by leakage current of PMOS access transistor and thus load transistor will be OFF.

**b) When data is logic ‘1’:** In this state, ST node pulled up to VDD-Vtn by NMOS access transistor, and therefore the drive transistor will be ON , and STB node will be pulled down to GND, thus load transistor will be ON which further pulls up ST node to VDD.

**3) Idle mode:** At the end of write operation, cell will go to idle mode and WL2 and BL are asserted to GND.

* **Read operation:**

When a read operation is issued the memory cell will go through the following steps.

**1) Bit-line Pre-charging:** For a read, BL pre-charged to GND and then floated. Since, in idle mode BL maintained at GND, this step doesn’t include any dynamic energy consumption.

**2) Word-line activation:** In this step WL2 asserted to VDD and two states can be considered.

**(a) Voltage of ST node is high:** When voltage of ST node is high, the voltage of BL pulled up to high voltage by NMOS access transistor. We refer to this voltage of BL as VBL-HIGH.

**(b) Voltage of ST node is low:** When voltage of ST node is low, the voltage of BL and ST node equalized. Since in this state, there is very small different between BL and ST node, power consumption is very small.

**3) Sensing:** The sense amplifier is enabled to read data on BL.

**4) Idle mode:** At the end of read operation, cell will go to idle mode and WL2 and BL are asserted to GND respectively.

**3.2 DESIGN CONSIDERATIONS**

To determine the sizes of the transistors in a SRAM cell and to ensure proper read and write operations, a number of design criteria must be taken into consideration. Design considerations for 6T and 4T SRAM cells are discussed as follows.

**3.2.1 DESIGN OF 6T SRAM CELL**

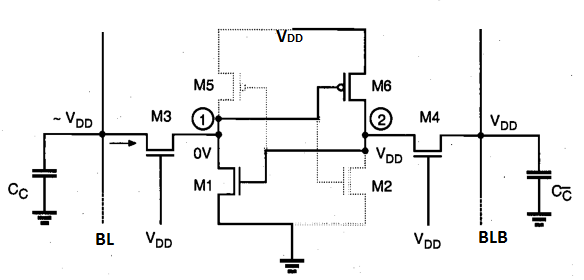
The two basic requirements which dictate the (W/L) ratios in a typical CMOS SRAM cell are:

(a) The data-read operation should not destroy the stored information in the SRAM cell.

(b) The cell should allow modification of the stored information during the data-write phase.

**Read “0" operation:**

Consider the data read operation first, assuming that logic "0" is stored in the cell. The voltage levels in the CMOS SRAM cell at the beginning of the "read" operation are depicted in Fig (3.1). Here, the transistors M2 and M5 are turned OFF, while the transistors M1 and M6 operate in the linear the linear mode. Thus, the internal node voltages are V1=0 and V2=VDD before the cell access (or pass) transistors M3 and M4 are turned on. The active transistors at the beginning of the data-read operation are highlighted in Fig (3.3).



**Fig (3.3) Voltage levels at the beginning of the read “0” operation**

After the pass transistors M3 and M4 are turned on by the row selection circuitry, the voltage level of column C̅ will not show any significant variation since no current will flow through M4. On the other half of the cell, however, M3and Ml will conduct a nonzero current and the voltage level of column C will begin to drop slightly. The column capacitance Cc is typically very large; therefore, the amount of decrease in the column voltage is limited to a few hundred mill volts during the read phase. The data- read circuitry is responsible for detecting this small voltage drop and amplifying it as a stored "0". While M1 and M3 are slowly discharging the column capacitance, the node voltage V1, will increase from its initial value of 0 V. Especially if the (W/L) ratio of the access transistor M3 is large compared to the (W/L) ratio of M1, the node voltage V1 may exceed the threshold voltage of M2 during this process, forcing an unintended change of the stored state.

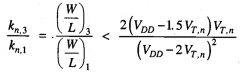
The key design issue for the data-read operation is then to guarantee that the voltage V1, does not exceed the threshold voltage of M2, so that the transistor M2 remains turned off during the read phase, i.e.

C:\Users\nikhila\Desktop\6tr1.PNG **Eq (3.1)**

We can assume that after the access transistors are turned ON the column voltage Vc remains approximately equal to VDD. Hence, M3 operates in saturation while M1 operates in the linear region.

C:\Users\nikhila\Desktop\eqre2.PNG **Eq (3.2)**

Combining this equation with Eq(4.1) results in:

 **Eq (3.3)**

The upper limit of the aspect ratio found above is actually more conservative, since a portion of the drain current of M3 will also be used to charge-up the parasitic node capacitance of node1.

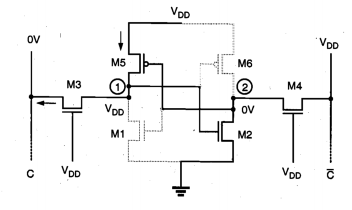
To summarize, the transistor M2 will remain in cut-off during the read "0" operation if Eq (3.3) is satisfied. A symmetrical condition also dictates the aspect ratios of M2 and M4.

Now substituting VDD = 0.5V and VT,n = 0.2V ; standard values in 50nm technology in “Microwind”, we get Eq (3.3) as W3 ≤ 8(W1) and similarly we obtain W4 ≤ 8( W2). From the layout W1 = W2 = 9 lambda, hence we obtain

W3 ≤ 72 lambda and W4 ≤ 72 lambda.

**Write “0” when “1” is stored:**

Now consider the write "0” operation, assuming that a logic "1" is stored in the SRAM cell initially. Fig (3.4) shows the voltage levels in the CMOS SRAM cell at the beginning of the data-write operation. The transistors M1 and M6 are turned OFF, while the transistors M2 and M5 operate in the linear mode. Thus, the internal node voltages are V1 = VDD and V2 = 0 V before the cell access (or pass) transistors M3 and M4 are turned ON.

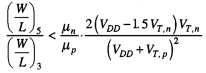


**Fig (3.4) Voltage levels at the beginning of the write "0" operation.**

The column voltage VC is forced to logic "0" level by the data-write circuitry; thus, we may assume that VC is approximately equal to 0 V. Once the pass transistors M3 and M4 are turned ON by the row selection circuitry, we expect that the node voltage V2remains below the threshold voltage of Ml, since M2 and M4 are designed according to Eq (3.3). Consequently, the voltage level at node (2) would not be sufficient to turn ON Ml. To change the stored information, i.e., to force V1 to 0 V and V2to VDD, the node voltage V1 must be reduced below the threshold voltage of M2, so that M2 turns OFF first. When V1 = VTn the transistor M3 operates in the linear region while M5 operates in saturation.

C:\Users\nikhila\Desktop\eq wr1.PNG**Eq (3.4)**

C:\Users\nikhila\Desktop\eqwr2.PNG**Eq (3.5)**

**Eq (3.6)**

To summarize, the transistor M2 will be forced into cut-off mode during the write"0" operation if Eq (3.6) is satisfied. This will guarantee that Ml subsequently turns ON, changing the stored information. Note that a symmetrical condition also dictates the aspect ratios of M6 and M4.

Now substituting VDD = 0.5V and VT,n = 0.2V ; standard values in 50nm technology in “Microwind” and taking µn/µp = 2.5 we get Eq (3.6) as

W3 > 2.45 (W5) and similarly we obtain W4>2.45 (W6).

From the layout W5 = W6 = 5λ, hence we obtain W3 > 12.25λ and W4> 12.25λ.

where 2λ= 50nm.

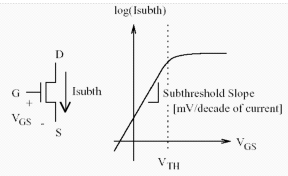
**3.2.2 DESIGN OF 4T SRAM CELL**

Write and read operations are performed using BL. In case of storing “0”, for data retention without refresh cycle, Eq (2.1) and Eq (2.2) must be satisfied. Sub-threshold leakage currents of access transistors are responsible for retaining the “0” value. Sub-threshold leakage currents and gate current of load and drive transistors have to be minimized.

High leakage current in deep-sub micrometer regimes is becoming a significant contributor to power dissipation of CMOS circuits as threshold voltage, channel length, and gate oxide thickness are reduced. Consequently, the identification and modeling of different leakage components is very important for estimation and reduction of leakage power, especially for low-power applications. Various transistor intrinsic leakage mechanisms are explained as follows.

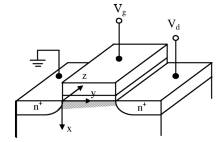
**Sub-threshold leakage currents:**

Sub-threshold or weak inversion conduction current between source and drain in NMOS transistor occurs when gate voltage is below VT. The weak inversion region is seen in Fig (3.5) as the linear region of the curve (semilog plot).



**Fig (3.5) Sub-threshold leakage in NMOS transistor**

In the weak inversion, the minority carrier concentration is small, but not zero. Fig (3.6) shows the variation of minority carrier concentration along the length of the channel for n-channel MOSFET biased in the weak inversion region. Let us consider that the source of the n-channel MOSFET is grounded. Gate voltage, VG<VT and the drain to source voltage VDS > 0.1V. For such weak inversion condition, VDS drops almost entirely across the reverse-biased substrate-drain p-n junction. As a result, the variation of the electrostatic potential ØS at the semiconductor surface along the channel (the y axis) is small. The component of the electric field vector E(Ey) being equal to, is also small. With both the number of mobile carriers and the longitudinal electric field small, the drift component of the sub-threshold drain-to-source current is negligible. Therefore, unlike the strong inversion region in which the drift current dominates, the sub-threshold conduction is dominated by the diffusion current. The carriers move by diffusion along the surface similar to charge transport across the base of bipolar transistors.



**Fig (3.6) Variation of minority carrier concentration**

The exponential relation between driving voltage on the gate and the drain current is a straight line in a semilog plot of ID versus VG. Weak inversion typically dominates modern device off-state leakage due to the low VT. The weak inversion current can be expressed as:

C:\Users\nikhila\Desktop\subthresheq1.PNGC:\Users\nikhila\Desktop\subthresheq1cont.PNG **Eq (3.7)**

**C:\Users\nikhila\Desktop\subthresheq2.PNG Eq (3.8)**

Where Vth is the threshold voltage, and *V*T = kT/q is the thermal voltage. Cox is the gate oxide capacitance; µo is the zero bias mobility; and m is the sub-threshold swing coefficient (also called body effect coefficient). Wdm is the maximum depletion layer width, and tox is the gate oxide thickness. Cdm is the capacitance of the depletion layer.

**Gate oxide tunneling:**

Reduction of gate oxide thickness results in an increase in the field across the oxide. The high electric field coupled with low oxide thickness results in tunneling of electrons from substrate to gate and also from gate to substrate through the gate oxide, resulting in the gate oxide tunneling current. When a positive bias is applied to the gate, due to the small oxide thickness, which results in a small width of the potential barrier, the electrons at the strongly inverted surface can tunnel into or through the SiO2 layer and hence give rise to the gate current. On the other hand, if a negative gate bias is applied, electrons from the poly-silicon can tunnel into or through the oxide layer and give rise to the gate current.

**Reducing leakage currents of load and drive transistors:**

Sub-threshold leakage currents come into picture when VGS < VT. When “0” is stored in the cell, load and drive transistors are OFF. Since we have increased VSS value to 0.04V, when load transistor is in OFF state, its VGS value becomes -0.04V which reduces the sub threshold leakage current contribution of load transistor. Here, VDD value is decreased to 0.45V so when drive transistor in OFF state its VSG value decreases, which reduces the sub threshold leakage current contribution of drive transistor. As VGS and VSG values of load and drive transistors have reduced, gate tunneling effect reduces. All these conditions are desirable as per Eq (2.1) and Eq (2.2). To reduce the leakage currents in sub-threshold region we have to reduce VTH which can be achieved by increasing the doping. In the tool, since the doping concentrations cannot be changed, we consider body biasing. The change in body biasing affects VTH which is called “Body effect”.

**Body Effect:**

The body effect describes the changes in the threshold voltage due to change in the source-bulk voltage, VSB. Since the body influences the threshold voltage, it can be thought of as a second gate, and is sometimes referred to as the "back gate". The body effect is also called the "back-gate effect". If the positive voltage on the substrate is increased, the minority holes are repelled and the channel can be formed easily, thereby reducing the threshold value.

For an enhancement mode n-channel MOSFET, body effect upon threshold voltage is computed using the following equation.

V_{TN} = V_{TO} + \gamma ( \sqrt{ | {V_{SB} + 2\phi_{F} | } } - \sqrt{ | 2\phi_{F} | } ) **Eq (3.9)**

Where, VTN is the threshold voltage when substrate bias is present

 VSB is the source-to-body substrate bias

 ØF is the surface potential

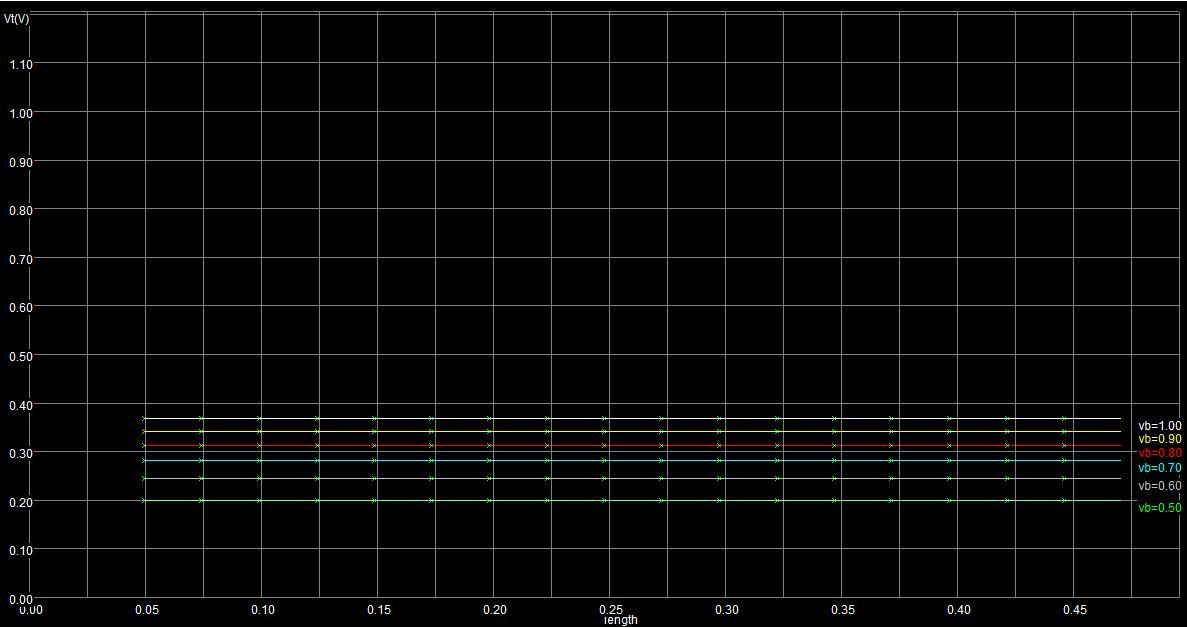
 VTO is the threshold voltage for zero substrate bias.

\gamma = (t_{ox}/\epsilon_{ox})\sqrt{2q\epsilon_{si}N_A} **Eq (3.10)**

γ is the body effect parameter, t_{ox} is oxide thickness, \epsilon_{ox} is oxide permittivity,\epsilon_{si} is the permittivity of silicon, NA is the doping concentration and q is the charge of an electron.

**Changes in VTH with change in N-well biasing:**

From the graphs shown below we can observe the changes in VTH with change in Nwell biasing as follows.

**Fig (3.7) Change in VTH with N-well biasing**

**Load transistor body biasing to read “1” delay trade off:**

To reduce the leakage current of load transistor we have increased VSB in order to increase VTH. But when the threshold voltage is increased, current IDS in saturation region will decrease.

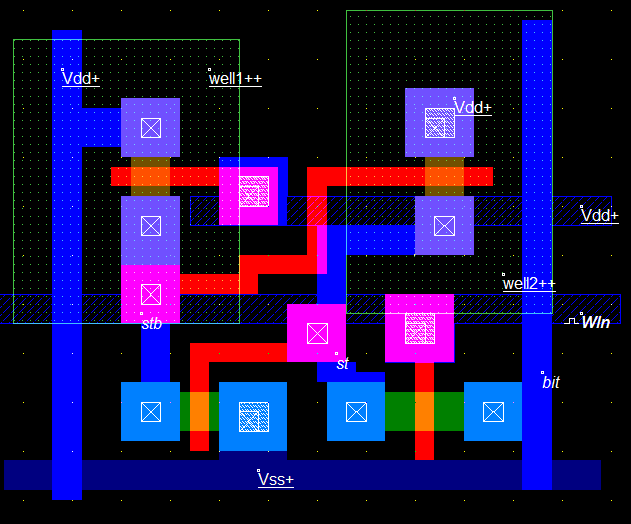
Current equation in saturation region is given by

I­ds = **Eq (3.11)**

From the above equation we can observe then when VT is increased current is decreased with power of two. If the bit line capacitance is large, more current is required to pull the bit-line to the value that is stored inside the cell. Since the current supplied by the load transistor is less more time is required to provide sufficient charge to overcome the capacitance.

But if the threshold voltage is decreased, it results in increase of leakage currents in turn increasing the delay in writing “0” and increasing the power consumed in write “0” and hold “0” condition.

So there exists a trade-off between the body biasing voltage and delay during read“1”.

****

**Fig (3.8) 4T layout with two N-wells**

**Values used:** VDD= 0.45V VSS = 0.04V

VTN0 = 0.2V VTP0 = 0.2V

VTN = 0.16V VTP,ACCESS = 0.24V

VTP,DRIVE = 0.28V

VNWELL, ACCESS = 0.6V

VNWELL, DRIVE = 0.7V

VWLN = 0.5 and 0V (ON and OFF)

VWLP = 0.3V

From the layout,

WNMOS,ACCESS= 6λ WLOAD = 6λ

WPMOS,ACCESS= 4λ WDRIVE = 4λ

Where 2λ= 50nm

**3.3 SUMMARY**

In this chapter working and design considerations of 4T and 6T SRAMs are discussed. For 6T SRAM, widths of transistors required to attain maximum stability of the cell are calculated. For 4T SRAM, different conditions for proper data storage are considered and parameters like width of the transistors, threshold voltages of transistors, supply voltages, pre-charge values and body biasing voltages are determined.