

SRAM analysis

Made of 6 transistors in which 4 are cross coupled transistors and there are 2 access transistors.

Terminology :-
 Pull-up (PU)
 Pull-down (PD)
 Access (Ax)

The cell must be symmetric that is each of the two PD, PD, Ax transistors must be sized the same way.

In particular, we will look at Read analysis, Write analysis & sense amplifier functionality.

READ

PRE = 1 (that is Pre charge transistors are off)

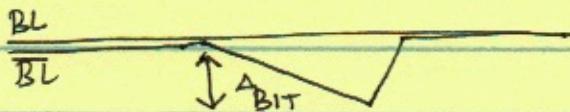
WL = 1 (the particular word line that is accessed must be turned on.)

The Bit lines (BL , \bar{BL}) are pre-charged to V_{DD} .

One of the two bit-lines will slowly begin to discharge depending on what value the SRAM cell is storing.

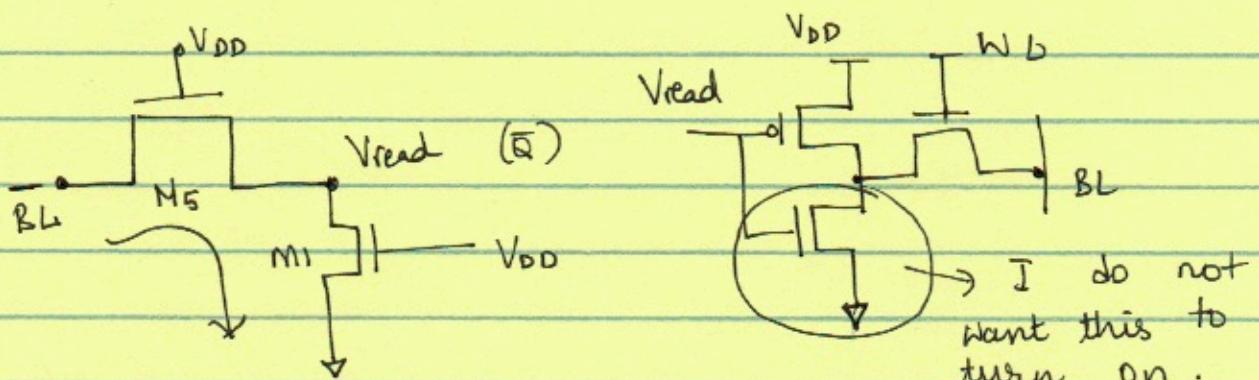
If $Q=1$ then \bar{BL} begins to discharge
 If $Q=0$ then BL begins to discharge

We want the BL or \bar{BL} to discharge by an amount Δ_{BIT} that the sense amplifier can detect.



In the above case, BL discharges by Δ_{BIT} indicating that the cell was storing a "1".

A key idea during Read operation is that while reading the cell value must NOT change. If you do not design the cell properly, you will enter into a problem called DESTRUCTIVE READ, which changes the stored value in the cell.



$$\text{So } V_{read} < V_{trip}$$

During the read operation, the voltage at Q (in the above example) will increase slightly by an amount V_{read} . If V_{read} is less than the switching threshold of the inverter then we are safe.

If $\bar{Q} = 0$ then $Q = 1$ and I do not want M₂ to turn on. If M₂ turned on, then Q will go to zero causing destructive read.

(A) Read noise margin :- $V_{trip} - V_{read}$

$V_{read} :- ?$

To find V_{read} , notice M₅ is in saturation and M₁ is in linear mode

$$0.5 \beta_{M5} (V_{DD} - V_{read} - V_{TN})^2 = \beta_{M1} \left(V_{DD} - V_{TN} - 0.5V_{read} \right) \frac{V_{read}}{\beta_{M1}}$$

This is a quadratic equation in V_{read} .

Solve it to get :-

$$V_{read} = (V_{DD} - V_{th}) \left[1 - \sqrt{1 - \frac{1}{1 + \frac{\beta_{M1}}{\beta_{M5}} \text{Ratio pd-ax}}} \right]$$

Where $\text{Ratio pd-ax} = \frac{\beta_{M1}}{\beta_{M5}}$

Important

If $\text{Ratio pd-ax} \uparrow$	$V_{read} \downarrow$
$\text{Ratio pd-ax} \downarrow$	$V_{read} \uparrow$

Also note that read voltage does not depend on the size of the pull-up device (M_3, M_4) in the SRAM cell.

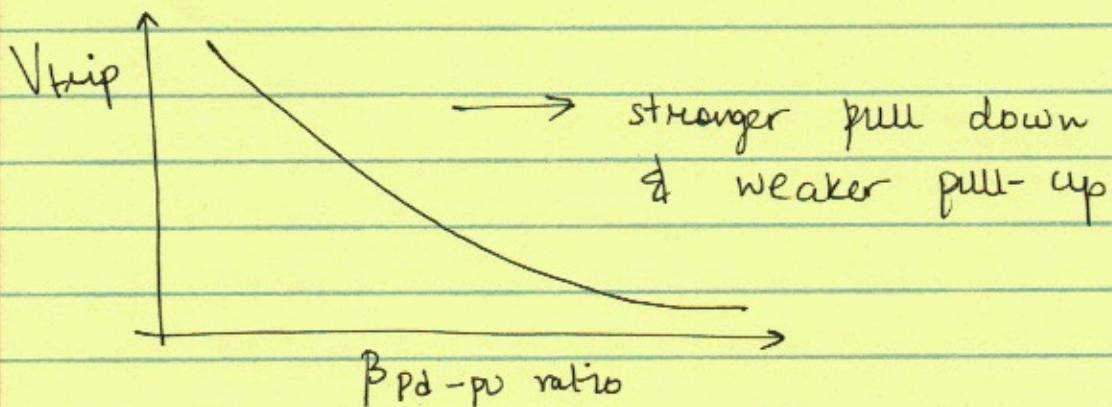
The pull-up device affects the trip voltage

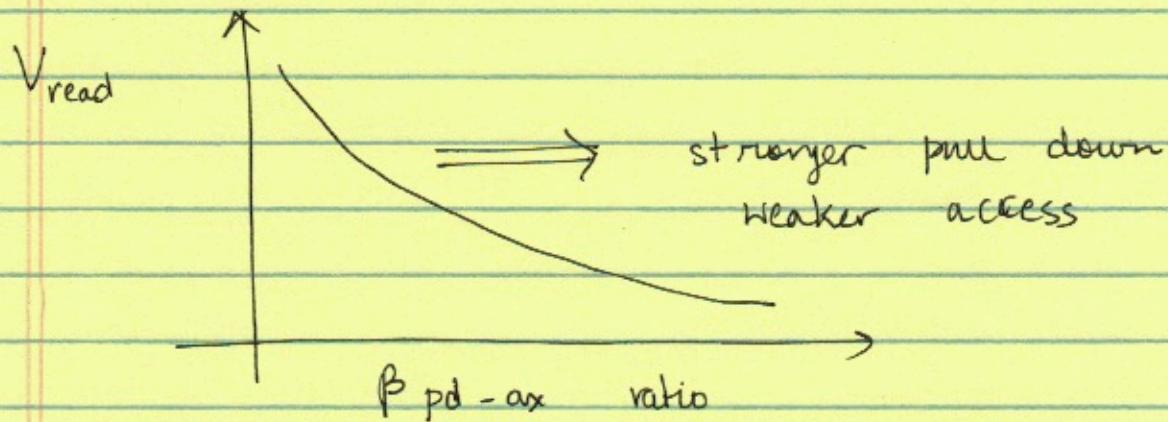
$$V_{trip} = \frac{V_{DD} - |V_T| + \sqrt{\beta_{ratio\ pd-pup} V_{thN}}}{1 + \sqrt{\beta_{ratio\ pd-pup}}}$$

As pull-down becomes stronger $V_{trip} \downarrow$

As pull-up becomes stronger $V_{trip} \uparrow$

Bad thing is if we want read voltage to be small, we want to make pull-down strong. But making pull-down strong, you end up making V_{trip} also small.





(B)

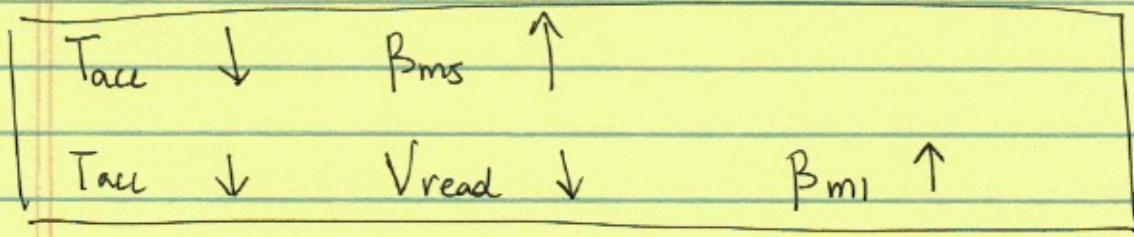
Read access time

$$T_{\text{access}} = C_{\text{bit}} \frac{\Delta_{\text{bit}}}{I_{\text{read}}}$$

I_{read} \Rightarrow current that flows through the access & pull down transistors

$$I_{\text{read}} = \frac{\beta_{m5}}{2} (V_{DD} - V_{\text{read}} - V_{TN})^2 \Rightarrow M5 \text{ is in saturation.}$$

$$T_{\text{acc}} = \frac{C_{\text{bit}} \Delta_{\text{bit}}}{0.5 \beta_{m5} (V_{DD} - V_{\text{read}} - V_{TN})^2}$$



Note :- Pull-up device does not play a role in deciding the speed of read access.

To Summarize

- (i) V_{TRIP} depends on the pull-up and pull-down device size ratios
- (ii) V_{read} depends on the access and pull-down device size ratios
- (iii) T_{acc} depends on the size of access transistor and V_{read} .

$$\text{FAST ACCESS} \implies \beta_{m5} \uparrow \quad \beta_{m1} \uparrow$$

$$V_{read} \downarrow \implies \frac{\beta_{m1}}{\beta_{m5}} \uparrow$$

$$V_{trip} \uparrow \implies \frac{\beta_{m3}}{\beta_{m4}} \downarrow$$

$$\text{FAST ACCESS} \implies \text{strong access} \& \text{strong pull down}$$

contradictory

$$V_{read} \downarrow \implies \text{strong pull down} \& \text{weak access}$$

contradictory

$$V_{trip} \uparrow \implies \text{strong pull up} \& \text{weak pull down}$$

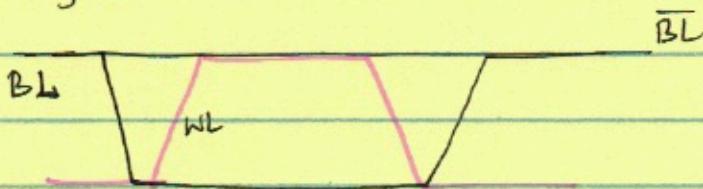
We clearly see a contradiction in the requirement of the device sizes. Typically, V_{read} is more sensitive to the size of the pull down device, compared to V_{trip} .

∴ typically, we make pull-down stronger to lower V_{read} but we can't make it too strong because even though V_{trip} is less sensitive to the size of pull-down, it is nonetheless sensitive. And we do not want to lower V_{trip} too much. Remember, ultimately we care about

$$V_{trip} - V_{read}$$

WRITE OPERATION

When we want to write into the cell, say initially the cell is storing "1" and we want to write "0". Then we make $\overline{BL} = 0$ before the particular WL goes high.



\overline{BL} stays high

BL makes "1" to "0" transition since we're trying to write "0" into the cell.

The key idea is node "Q" in the cell will discharge through the access transistor since BL is pull down to 0.

When Q is discharging through the access, then the pull-up transistor is pushing in some charging current & tries to prevent Q from discharging all the way to 0. \therefore we need to make sure through proper cell sizing that the node storing "1" can be pulled below the switching threshold of the other inverter.

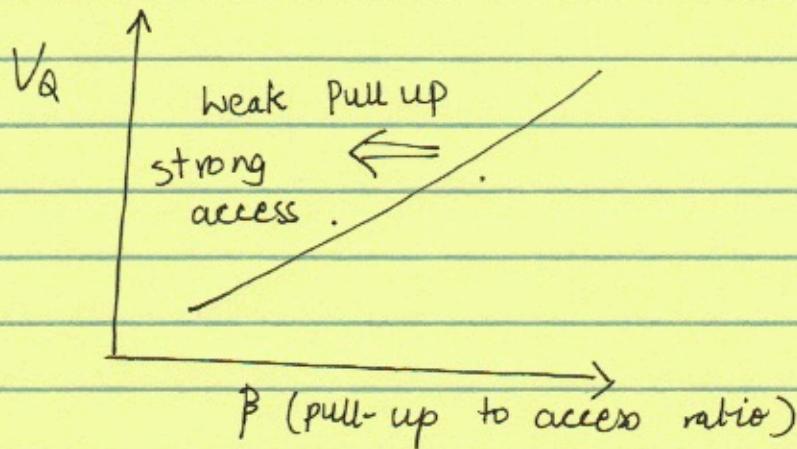
So for write operation we analyze the discharging of node "Q" to ground.

Since we want V_Q to go below the switching threshold of the inverter, we assume V_Q is very low. So for low V_Q , M_6 (access) will be in linear mode while M_4 (pull-up) will be in satn.

$$0.5 \beta_{m4} (V_{DD} - |V_{TP}|)^2 = \beta_{m6} [(V_{DD} - V_{TN}) - 0.5V_Q]V_Q$$

WRITE Voltage $\leftarrow V_Q = (V_{DD} - V_T) \left(1 - \sqrt{1 - \frac{\beta_{pu}}{\beta_{acc}}} \right)$

$$V_T = |V_{TP}| = V_{TN}$$



For better writability we want weak pull up and strong access

Thumb rule

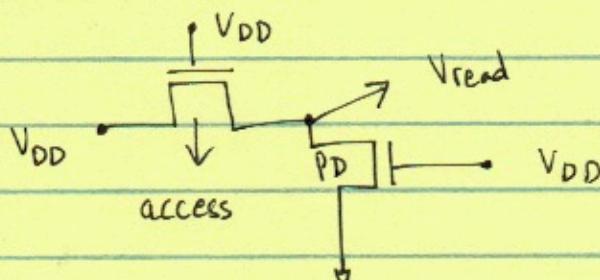
$$W_{pu} : W_{ac} : W_{pd} = 1 : 1.5 : 2$$

But you should carefully size the devices in each new technology. This can be a starting point in the SRAM design project but you would have to optimize this design to meet the target specs.

In your design project you have to specifically measure the read & write margins of the cell.

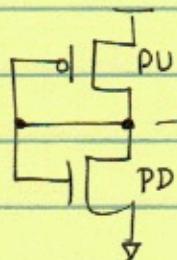
Here is how you will measure these margins

Read margin



Measure the DC point voltage V_{read} of the cell.

Then for V_{trip} do the following:-

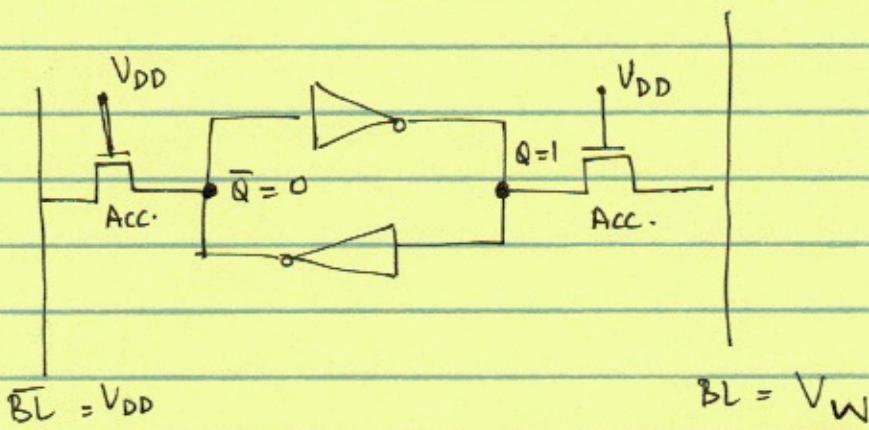


V_{trip} measure the DC point voltage by shorting the inverter input & output terminals.

Once you have measured V_{read} and V_{trip} in your Cadence simulations :-

Noise margin for read :- $\underline{V_{trip} - V_{read}}$

WRITE margin for read



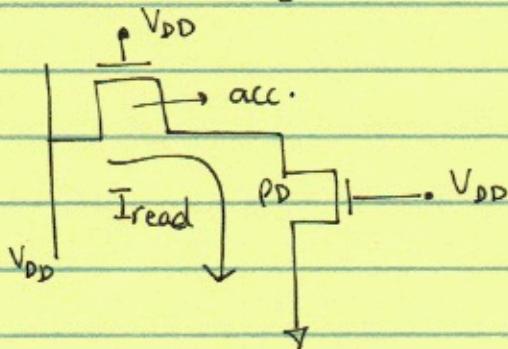
Let's say the cell is storing "1" and now you want to write "0" into the cell.

Ideally you will lower BL to 0.
 But to measure the write margin, increase BL slowly and find the highest voltage V_W at which you can reliably write into the cell.

Write margin :- V_W

Measurement of access time :-

Use the same circuit as we did for read voltage measurement.



Measure the DC current of the circuit.

$$T_{acc} = C_{bit} \frac{\Delta_{bit}}{I_{read}}$$

$$C_{bit} = 40 \text{ fF} \quad \text{y initial estimate.}$$

$$\Delta_{bit} = 10 \text{ mV}$$

SENSE AMPLIFIER

In the case of SRAM, the purpose of sense amplifier is to resolve the stored bit in the memory w/ small signal swings on the bit line.

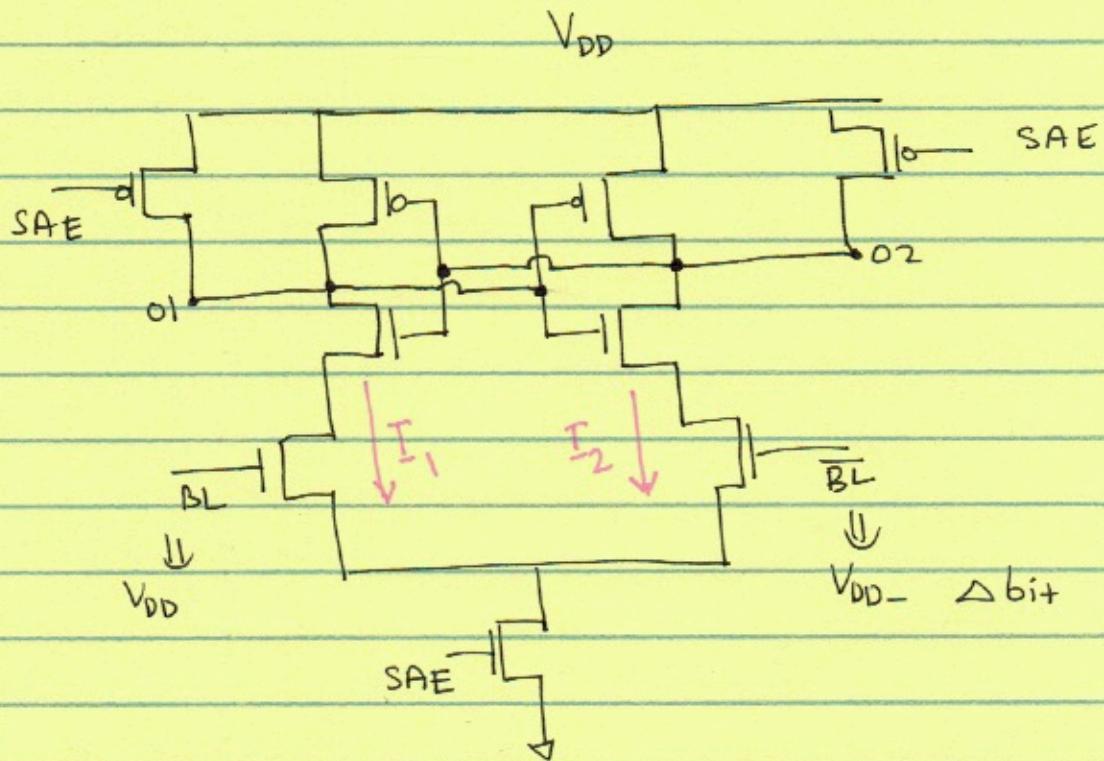
- Small signal swings on the bit line ~~come~~ help to reduce the delay & power dissipation.

Sense amplifiers are analog circuits by default.

Differential voltage Sensing

- Takes small signal differential input (bit line voltages) and amplifies it to large signal output.

In a current latch based sense amplifier differential input is converted to a differential output. That is, both the output & its complementary signal are generated.

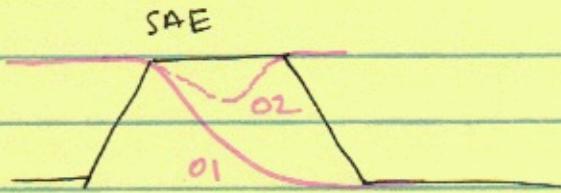


$SAE = 0 \rightarrow$ Sense amplifier is not really sampling BL and \overline{BL} .
 Nodes O_1 and O_2 are both precharged to V_{DD} .

Only when $SAE = H \wedge H \rightarrow$ sense amplifier is able to evaluate the difference in signal BL and \overline{BL} .

When \overline{BL} discharges because the cell was storing a "0", $I_1 > I_2$

This creates a positive feedback and O_1 begins to discharge faster than O_2 .

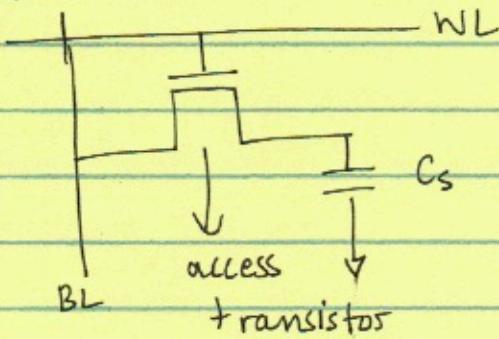


When the voltage difference b/w the nodes o1 & o2 increases, the cross-coupled inverter pair amplifies it. A latch at the output of the sense amplifier generates rail-to-rail static complementary output signals o1 & o2.

Note:- Since this sense amplifier works on the difference in currents I_1 & I_2 , even a small amount of mismatch leads to design challenges.

To account for the transistor mismatch that comes from the fabrication process, a minimum bit differential is required for sense amplifier to work properly. This minimum voltage differential is also called the sense amplifier offset.

1+ Dynamic Random Access Memory (DRAM)

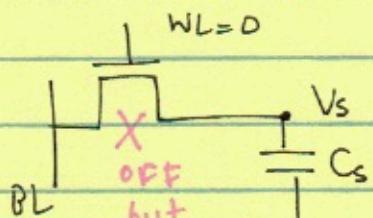


Important features

- (a) Value of the bit is stored as charge on the capacitor C_s .
- (b) Single-ended memory, as only the bit is available but not its complement.
- (c) Sense amplifier is a must for every bit line for correct functionality → **Important**
↳ Because of the charge re-distribution read out
- (d) Has a maximum retention time. After that the cell must be refreshed or it will lose the data.
- (e) When writing "1" into the cell, a threshold voltage is lost. This can be circumvented by bootstrapping the WL to a voltage higher than V_{DD} .
- (f) Single-ended sense-amplifier is just an inverter, but it is susceptible to noise. So single-ended o/p is converted to differential output.

Issue of charge leakage in DRAM

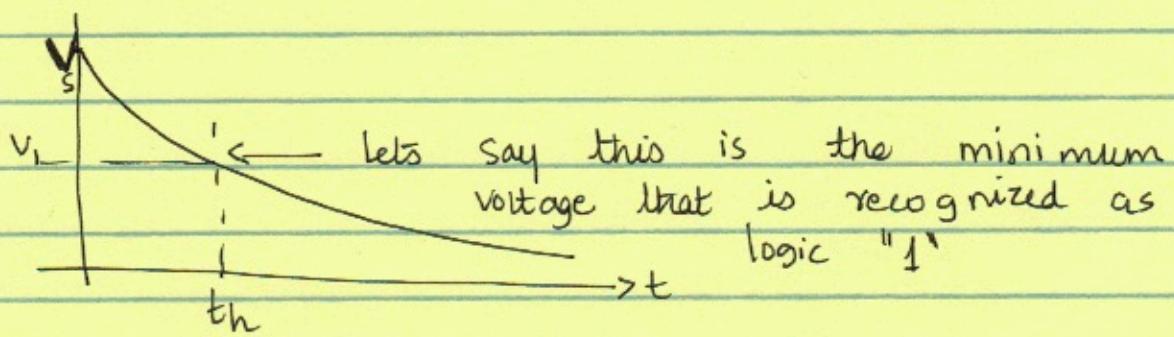
Lets say the cell is storing "1" and the WL is off.



but there will be a leakage current I_L (same as sub threshold .. and junction leakage - - -)

to store "1" the charge has voltage V_s

But because $WL = 0$, there will be leakage as V_s now is in a high-impedance state. The charge stored on C_s will begin to leak.

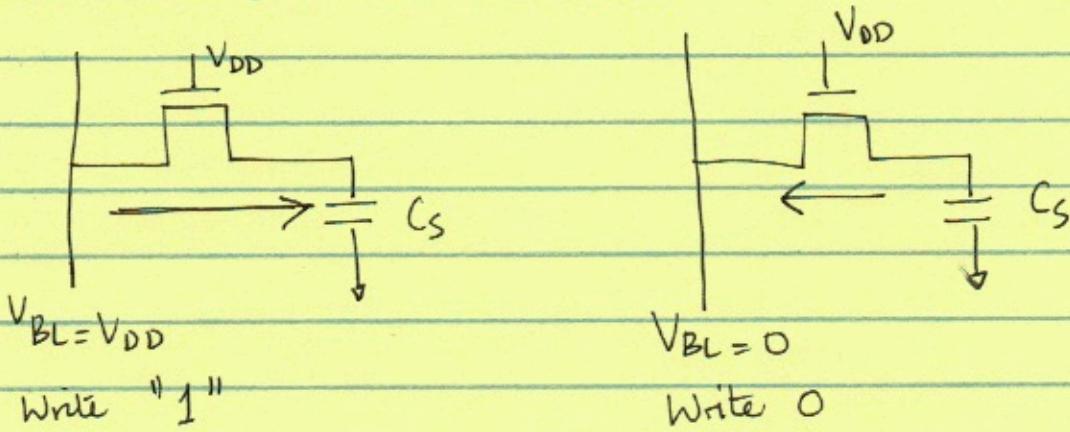


$$t_h = \frac{C_s}{I_L} \Delta V_s = \frac{C_s}{I_L} (V_s - V_L)$$

After a time of t_h , the DRAM cell will need to be refreshed.

WRITE analysis

The Data value is placed on the BL and the WL is raised. Depending on the data value on the BL, the cell capacitance is either charged or discharged.



For writing "1", the capacitance is charged up.
For writing "0", the capacitance is discharged.

Note, the maximum voltage to which C_S can be charged is $(V_{DD} - V_{TN})$ because of threshold voltage drop.

READ OPERATION

BL is precharged to V_{PRE} .

V_{PRE} is typically $\frac{V_{DD}}{2}$.

When WL goes high then depending on the value of the data stored in the cell, the voltage on BL can go above V_{PRE} or go below V_{PRE} .

If the cell is storing "1" then V_{PRE} increases.

If the cell is storing "0" then V_{PRE} decreases.

The increase or decrease in V_{PRE} is sensed by a sense amplifier to figure out if the cell was storing "1" or "0".

Since V_{BL} voltage changes during the read operation, it also means that the charge that was initially stored on C_s changes.

Since the total charge is conserved, if $V_{BL} \uparrow$ it means charge on $C_s \downarrow$.

If $V_{BL} \downarrow$ it means charge on $C_s \uparrow$.

Ex: Q_0

READ OPERATION IN DRAM IS ALWAYS DESTRUCTIVE. That is, it changes the stored charge

The change in the bit line voltage is given as :-

$$\Delta V = (V_{B\text{IT}} - V_{\text{pre}}) \left(\frac{C_s}{C_s + C_{BL}} \right)$$

C_s = storage cap

C_{BL} = Bit line capacitance

↳ charge transfer ratio of the DRAM cell.

$$C_s \ll C_{BL} \Rightarrow \Delta V \text{ is tiny.}$$

Let's say, we were reading "0"

$$V_{B\text{IT}} = 0 \quad \& \quad V_{\text{pre}} = V_{DD}/2$$

$$\Delta V = -\frac{V_{DD}}{2} \left(\frac{C_s}{C_s + C_{BL}} \right)$$

$$\text{If } C_s/C_{BL} = 0.05 \Rightarrow \Delta V \approx 62.5 \text{ mV for } \underline{V_{DD}=2.5V}$$

Let's say we were reading "1"

$$V_{B\text{IT}} = V_{DD} - V_{TN} \approx 2V$$

$$V_{\text{pre}} = V_{DD}/2 = 1.25V$$

$$\Delta V = (2V - 1.25) \left(\frac{1}{1+20} \right) \approx \underline{\underline{35 \text{ mV}}}$$

SENSE AMPLIFIER DESIGNS FOR DRAM

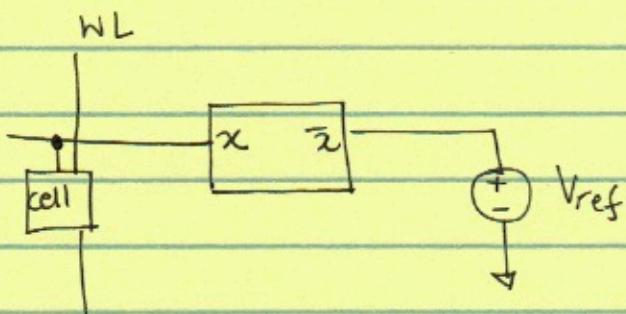
single-ended sensing

simply an inverter
with a specific
trip voltage.

Convert single ended
into differential
sensing scheme

compare the BL
voltage with
a reference voltage
using a differential
sense amp.

Differential
sensing



Creating the reference voltage

Memory array is divided into two halves,
with the differential amplifier placed in
the middle.

On each side a column of dummy cells
is added. These are 1T DRAM cells but
they serve as reference.

Approach is called open bit line architecture with
dummy cells.

- a) When EQ is raised both BLL and BLR are precharged to $V_{DD}/2$
- b) Enabling L and \bar{L} at the same time ensures that the dummy cell capacitances are charged to $V_{DD}/2$.
- c) One of the WL is enabled during the read operation.
- d) Let's say a cell in left hand side array is selected. This leads to a change in BLL.
- e) The appropriate voltage reference is created by simultaneously selecting the dummy cell in the other half of the memory cell by raising L.
- f) Now BLL and BLR are compared causing the sense latch to toggle.

Note raising the WL_0 and L simultaneously turns the capacitive coupling b/w BL and WL into a common mode signal which is effectively eliminated by the sense amp.