

Design of Low Leakage Low Voltage High Density 6T-SRAM Bit-Cell

¹Hariom Singh

Mtech Scholar
Bhabha University,Bhopal

² Satyarth Tiwari

Asst.Professor
Bhabha University,Bhopal

Abstract—A low retention voltage 6T static random access memory is analyzed in 40 nm CMOS technology and standard 6T-SRAM with high threshold voltage transistor for ultra-low leakage is used with area of used 6T-SRAM is $0.242 \mu\text{-m}^2$. In this work we reduced the rail to rail swing of 6T-SRAM cell by increasing the ground voltage which will decrease the leakage current at worst corner FF/125 of memory cell and also improve the stability of the memory cell at worst corner SF/ 40 and SF/125.

I. INTRODUCTION

In today's scenario, we are decreasing the size of transistor to make a compact design through which we can perform numerous of operations and also follow the moore's law also we can increase the density of chip through which reduce the cost of the fabrication but same Power density also increases which restrict the size of die[1]. Due to technology scaling some parameters of the device also fluctuates like gate oxide thickness, length and varied the some device characteristics. To reduce the power density of the chip by reducing the supply voltage and threshold voltage performance of memory is deterioration due to some unwanted parameters like sub threshold voltage, reverse diode leakage and power dissipation which decreases the performance of SRAM like stability and restrict the high speed applications[2]. But where less speed is required we can reduce the supply voltage. SRAM suffers from parametric variations due to smallest geometric size of SRAM cell.

Stability of the memory read is decided by static noise margin (SNM) or read margin (RM), retention stability is decided by the retention noise margin (RNM) and write stability of memory is decided by write margin (WM). Generally there are numbers of architectures of SRAM through we can use read and write operation of memory cell but most commonly used architecture is 6T-SRAM memory cell because it's nature of very high density with the help of this we can increase the size of memory on a single chip. Area and power consumption is two major part of system on chip (SoC) because 20 to 30 percentage of area is occupied by the SRAM and 40 to 50 percentage of total power of chip is consumed by the SRAM. But sometimes due to parametric variations there is a need to use 8T and 10-T SRAM memory. For low power consumption need to scale down the supply voltage but this will impact on the stability of the memory cell. Scaling the

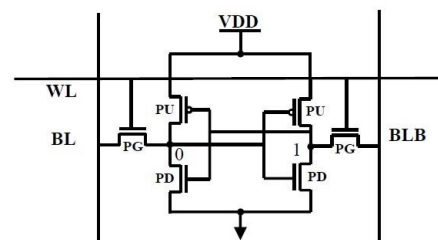


Fig. 1. 6T-SRAM Cell

supply voltage retention noise margin (RNM) static noise margin (SNM) and write margin (WM) is also decreases which can flip the data easily through which the stability of the memory cell decreases. To overcome the instability of memory we design some write and read assist technique through which we can work on low supply voltage without impact the performance of memory cell . For low leakage we generally use multiple threshold transistor, decrease the rail to rail swing of memory cell and dual power supply through which we can reduce the standby mode current and leakage power dissipation in 6T- SRAM cell[4].

As we are increasing the number of transistor on the system on chip (SoC), then power dissipation is considerable in both active as well as standby mode so to reduce the leakage we need to reduce the supply voltage as we discuss above. For standby mode or hold mode we need to reduce the leakage of memory cell by reducing the supply voltage of system on chip (SoC) in standby mode.

II. CONVENTIONAL 6T-SRAM CELL

Basic or conventional diagram of 6T-SRAM cell is shown in fig.1. In this memory-cell we used bi-stable latch which used to store or retain the data. PMOS is used to pull-up (PU) the voltage and NMOS is used to pull-down(PD) the voltage. Another two NMOS pass transistor is connected which is used to read and write the data into or out the bi-stable circuit. Pass gate (PG) is connected through bit-line (BL) and bit-line bar (BLB) and gate of pass transistor is connected to the word-line (WL) as shown in fig.1. In most cases SRAM cell is made with less size to increase the density of SoC and reduce the cost of chip. Because SRAM contains 20 to 30 percentage area of chip which is significantly very high.

III. BASIC ARCHITECTURE AND OPERATION OF 6T-SRAM CELL

Through 6T-SRAM cell three operation like memory read, memory write and retention mode can be performed. But the operation depends on the sizing of transistor because the same transistor is used for memory read and memory write operation. The operation mode of memory cell are as follows:-

- Hold or Retention Operation
- Read Operation

- Write Operation

A. Hold or Retention Mode of Operation

In this mode of operation, the memory cell retains the data up to some high value of noise is effected the data. The maximum value is required after that increasing the value or adding some noise flips the memory cell that value is called retention noise margin which describes the stability of retention mode. The circuit of retention mode is shown in Fig. 2. When the bit-line (BL) and bit-line bar (BLB) is floating or connected with high voltage and word-line (WL) is put on zero voltage, in that condition, the memory cell will store its data.

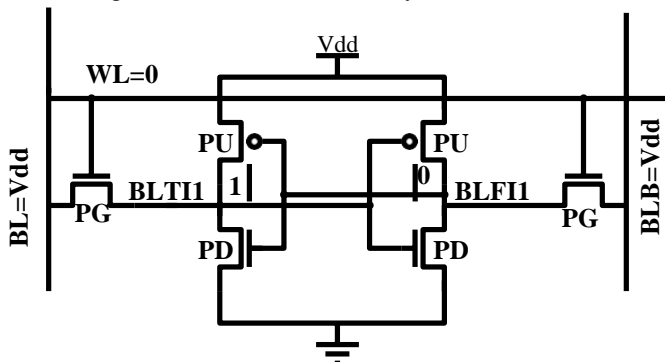


Fig. 2. Retention Mode

B. Read Operation

Fig. 3 describes the read operation of 6T-SRAM cell. Initially, charge the BL and BLB up to the Vdd. After charging the BL and BLB, put it floating and rise the voltage of word-line 0 to Vdd. Then both the pass gate (PG) is connected to the 0 and 1 node. So node 0 starts to charge through pass gate (PG) and BL starts to discharge, so I_{read} current flows from BL to node 0 through pass gate. BLB remains at high voltage Vdd (it discharges a little bit due to leakage but we can ignore that discharge). The discharge rate of bit-line depends on the capacitance of bit-line of memory cell. If a large number of columns are used, then bit-line capacitance will be high, so discharging of bit-line will be slow. The developed difference between bit-line and bit-line bar is sensed by the sense amplifier and evaluated as the output.

For proper read operation, the sizing of pass gate and pull-down transistor is very important. The size of pass gate should be less than the pull-down transistor. If we put a large pass gate, then high leakage current will flow through pass gate and

also bit-line load increases which slows down the access time of read operation. Cell stability is a critical parameter, so pass gate is typically weaker than pull-down transistor.

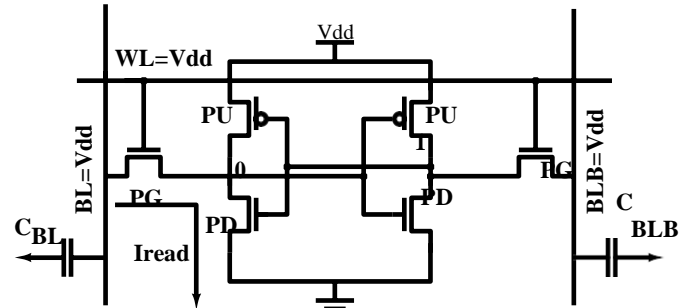


Fig. 3. Charging and Discharging of BL and BLB during Read Operation

C. Write Operation

In Fig. 4, the write operation of memory is shown. For write operation, pull the bit-line or bit-line bar at 0 V through word driver pulse which is controlled by write signal WR and WRB and now put word-line (WL) from low to high (0 to Vdd). Bit-line starts to discharge and try to pull down the node 1 via pass gate. At the same time, pull-up (PU) transistor tries to maintain the same value at node 1. In that case, current I₂ flows through pull-up and pass gate transistor. We make a powerful write driver and pass gate which avoid the condition of flipping the memory cell data.

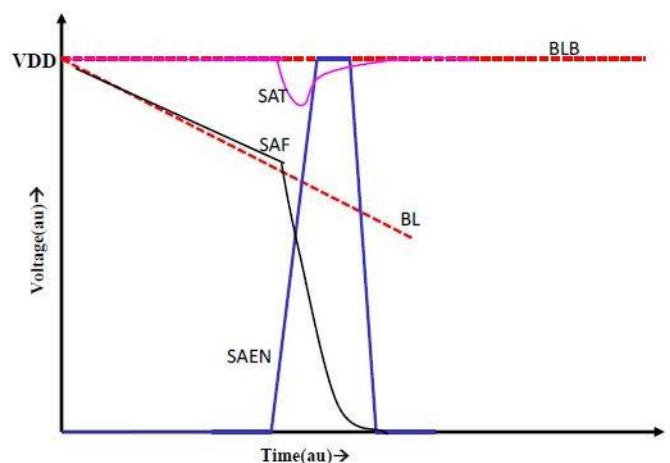


Fig. 4. Read Operation

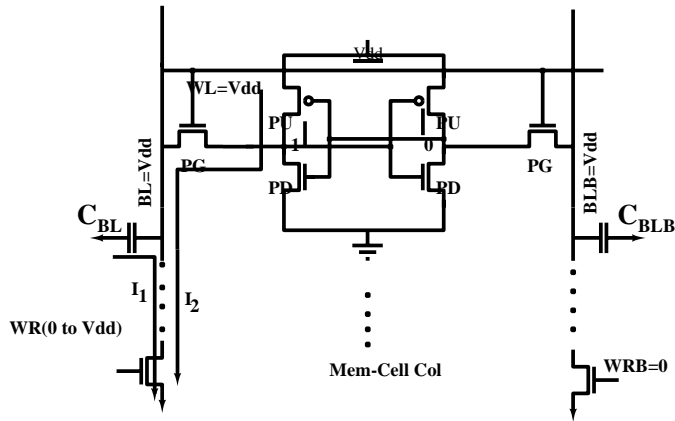


Fig. 5. Charging and Discharging of BL and BLB during Write Operation

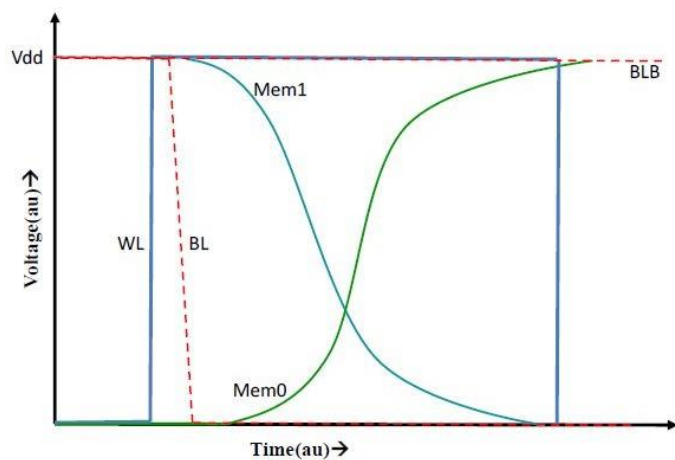


Fig. 6. Write Operation

IV. PREVIOUS CIRCUITS PROBLEM DESCRIPTION

A. Conventional Retention Mode

In this section we discussed about the conventional retention mode operation. From the circuit fig. we can analyzed easily that the stability of the circuit is very high and data is flipping at a very high retention voltage but from the fig.7. We can analyzed that the leakage current like sub-threshold current, bit-line leakage current, array leakage current and gate leakage is very high in standby mode. Through this circuit we can not reduce the leakage current in standby mode. To reduce the leakage current in standby mode we need to decrease the rail to rail swing in retention mode[8].

B. Diode connected Virtual Ground Retention Mode

As we discussed in previous section leakage current is very high so to reduce the leakage current in standby mode need to reduce the rail to rail swing in retention mode. As fig.11 shows the source bias replaced the ground voltage of this circuit. For source bias we used a diode connected NMOS which is controlled by the enable (EN) signal means when EN signal is high then it works in read or write mode

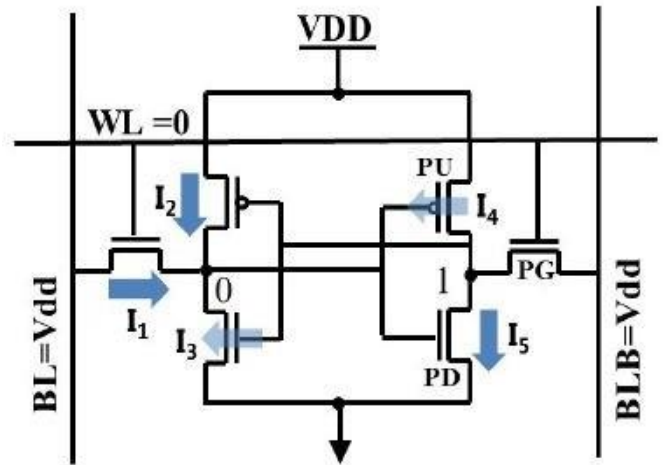


Fig. 7. Leakage Current

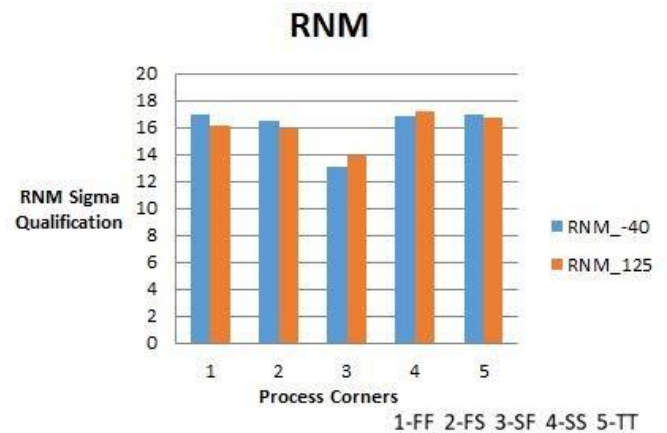


Fig. 8. Sigma Qualification with Process Corners

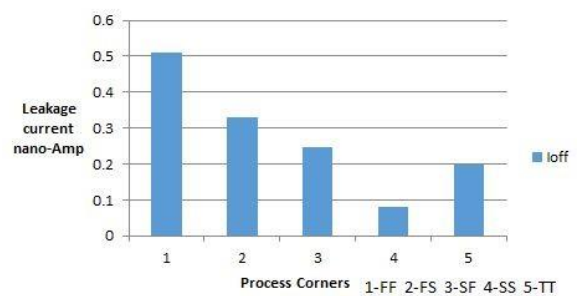


Fig. 9. Leakage Current

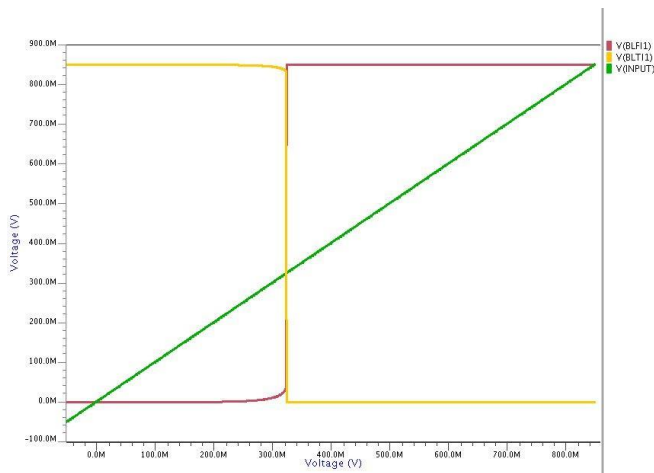


Fig. 10. Retention Voltage

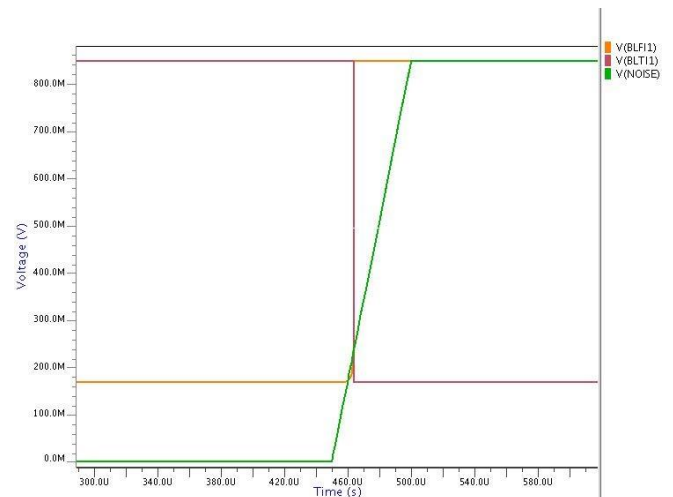


Fig. 12. Retention Noise Voltage

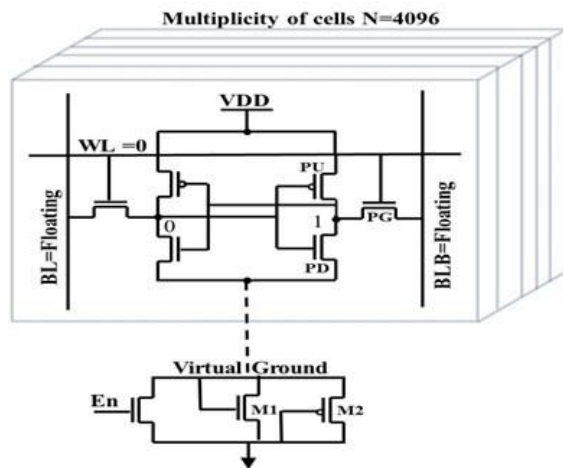


Fig. 11. 6T-SRAM Cell With Virtual Ground

but when EN signal is low then leakage current is passes through the diode connected NMOS and PMOS. Stability of the memory cell is decided by the rail to rail swing of the memory-cell[6].

$$V_{rail-to-rail} = V_{DD} - V_{virtual}$$

Lower the value of virtual the stability of the memory cell but same time less reduction in leakage current. So for proper virtual ground we used the combination of PMOS and NMOS for better cross corners because if we used only NMOS for virtual ground, then SF corner will be worst because for SF corner PMOS is fast and NMOS is slow. So for proper process corner we used the combination of PMOS and NMOS. Bit-line and bit-line bar will be floating to reduce the leakage current through that.

With the combination of PMOS and NMOS we can get the good process corners but the stability at SF corner is not good we can analyzed from the fig.13. Stability of the memory is decided by the sigma qualification. Sigma is the ratio of

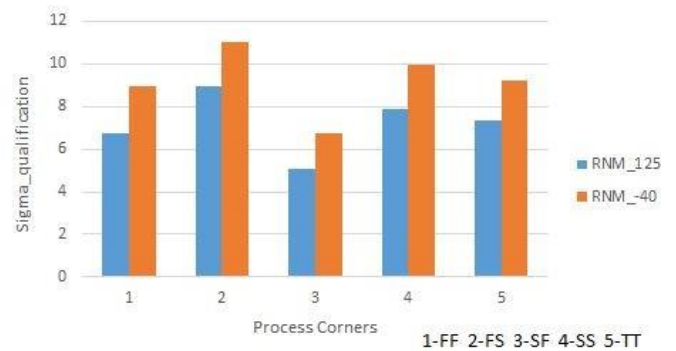


Fig. 13. RNM Sigma Qualification

RNM and standard deviation. As compare to the previous circuit leakage at standby mode is reduces but the stability at SF/-40 and SF/125 corner is not up-to-the mark. SF/-40 is most unstable point and at FF/125 highest leakage current in standby mode. This is due to the virtual ground voltage is controlled by the threshold voltage of the NMOS and PMOS. Threshold voltage of the device is increases as we decreases the temperature and as we increases the temperature threshold voltage is reduces. So the stability is less at less temperature compare to the high temperature.



Fig. 14. Leakage Current at Standby Mode

V. PROPOSED ARCHITECTURE

As we discussed the problem in source bias circuit, the stability at SF/40 is worst and the leakage current at FF/125 is maximum so to overcome these problems we designed a architecture in without diode connected NMOS and PMOS. Here we used a simple NMOS whose gate is controlled by some other circuit which is describe in the fig. 15. With the help of this circuit we overcome the problem of instability at SF/-40 and SF/125 corners.

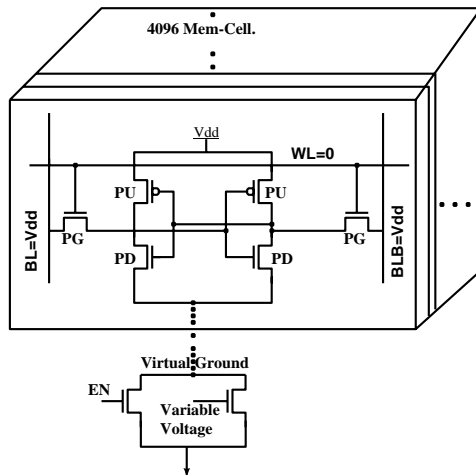


Fig. 15. 6T-SRAM with Virtual Ground and Variable Gate Voltage

For this variable gate voltage we used a another memory cell which is higher than the 4192 memory cell so that leakage will be more as compare to the 4192 memory cell, so gate voltage will be higher than the drain voltage of that memory cell. As we increased the gate voltage by some micro-volt as compare to the drain voltage the strength of NMOS is increases which will put the drain voltage towards ground and

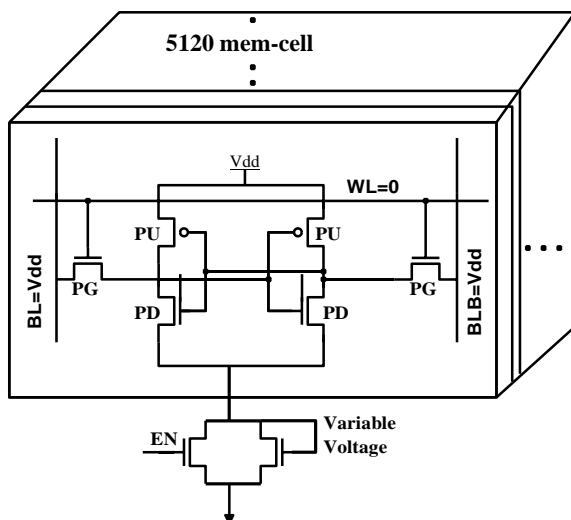


Fig. 16. Variable Gate Voltage

reduce the leakage and increased the stability of the memory cell in standby mode.

VI. RESULTS

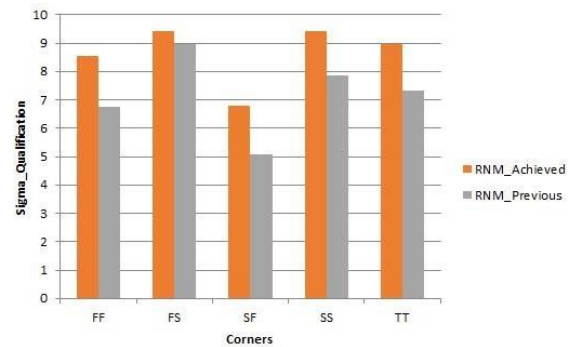


Fig. 17. Stability of Memory-Cell in terms of Sigma Qualification

The stability of the memory cell is improved by the proposed circuit. As we can observe from the fig. 16 the stability at SF corner is worst in previous case which is improve by large amount is new proposed circuit. Also the leakage current is reduced by 5 percent at FF corner as comparison to previous circuit.

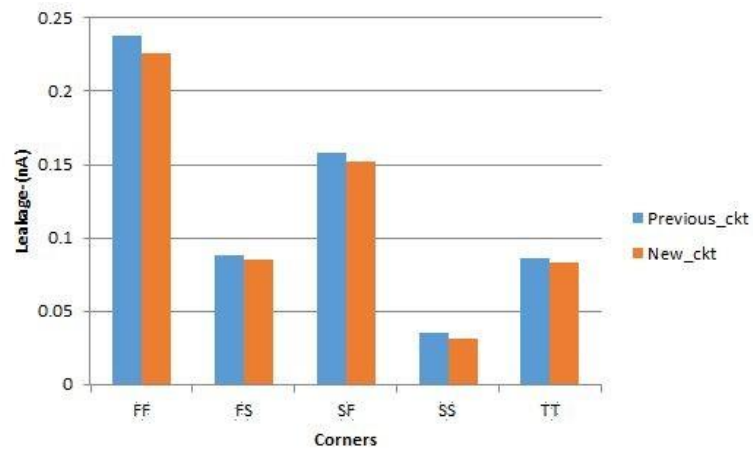


Fig. 18. Leakage of Memory-Cell

VII. MEMORY READ

Proper operation of memory read we need some read assist techniques because in 6T SRAM cell, we need same path for read and write operation. If we sized the memory transistor the either we can improve read or write of the memory. If we improve read margin of the memory, then we will lose the write margin and vice versa. So to overcome this problem we need some read and write assist technique for proper operation of memory read and write.

A. Read Assist Techniques

- Dual Power Supply
- Word-Line Under-Drive
- Reduced Bit-line capacitance or use shorter bit-line
- Sub-VDD bit-line pre-charge

Due to several advantages and disadvantages of the read assist we used word-line under-drive technique for read assist.

B. Dual Power Supply

In dual power supply SRAM, we used two power supply on for memory array power supply and second is for peripheral. In SRAM most of the leakage power of static power and dynamic power is dissipated by the peripheral of SRAM, so for low power consumption peripheral is put at less voltage as compare to the memory array voltage[5]. Word-line selector row decoder is also put on memory array voltage because row decoder drives the word-line of the memory array. As in figure shown the power supply of memory is VDDMA and power supply of bit-line is VDDMP. For less power consumption peripheral power supply is less than the power supply of memory arrays ($VDDMP > VDDMA$). But it is very difficult to generate dual power supply on SoC. So generally this scheme is less in use but if we need high speed memory than this scheme is best because in other scheme performance loss happened.

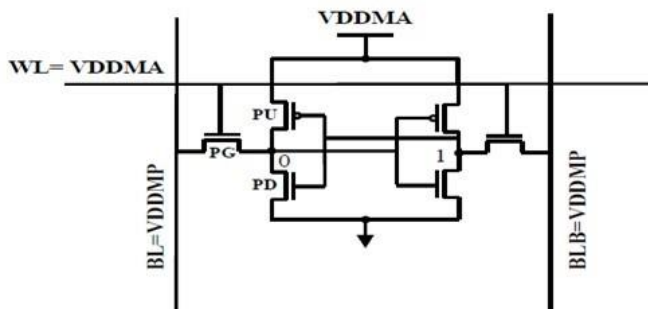


Fig. 19. Dual Power Supply SRAM

C. Word-Line Under-Drive

Word-line under-drive is most famous concept for better read margin. In this scheme we reduced the word-line voltage which is connected to the pass gate of the memory cell. So

decrement in word-line voltage it will decrease the strength of the pass gate and improve the read margin or static noise margin(SNM) of the 6T memory-cell. Word-line is decreases by some delta V voltage with the help of WLUD circuit corresponds to this changes in the circuit the static noise margin of memory is improved by god amount because we reduce the strength of pass gate and improve the strength of pull down transistor[5]. As for low leakage and less power consumption we need less power supply but as we will decrease the power supply, SNM also decreases so to improve SNM and dynamic noise margin (DNM) word-line under-drive(WLUD) scheme is very helpful. Through WLUD flipping time of memory cell is increases which reduces the data loss at low SNM and DNM. So with the help of WLUD probability of losing the data reduced.

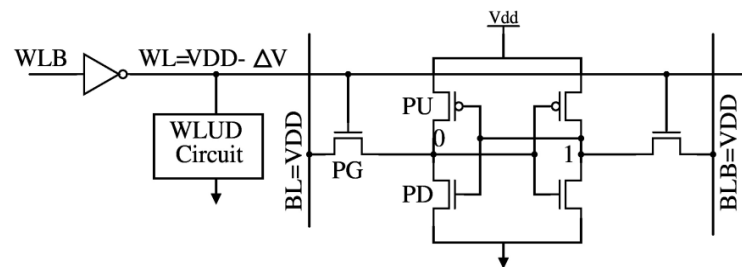


Fig. 20. 6T-SRAM Cell with WLUD

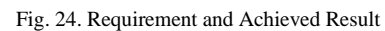
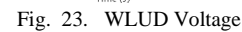
To increase the read stability of 6T SRAM memory cell proposed the architecture which is shown in fig. 20. In this figure we used a NOT gate as a word-line driver with high strength of PMOS which is followed by WL. 6T SRAM memory cell is a single port memory-cell it's means used a single path for both read and write operation. If we want to improve the read stability then needs to decrease the strength of pass gate and increase the strength of pull down transistor whereas for the write stability need to increase the strength of pass gate. So this is contradict between read and write stability in case of single port memory cell. To overcome this problem we designed an architecture to improve the read stability without impacting the write stability of memory-cell. In this architecture we decrease the strength of pass gate during the read operation of memory cell, through which will increase the flipping time within this time we can read the data without flipped it.

D. Proposed Architecture to Improve Read Stability

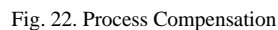
To under drive the word-line we used a PMOS but the worst corner corresponds to FS/125 at this corner high under driver is required and at SF corner no under driver is required. So to under driver the WL we need a process compensation circuit through which we can control the bleeder voltage of PMOS. For process compensation we designed one PMOS



So from figure we can looked that the voltage difference at FS and SF corner is high but still the value is high at FS corner so reduce this high value we require a sub-tractor which will subtract all corners values and by using NMOS we can put the drain value of PMOS at zero volt. So through PMOS sub-tractor the value at FS corner reaches to few mV value which is enough to under drive the value at FS corner and also used a temperature compensation circuit through one NMOS and one PMOS connected in diode fashion. The output of this circuit is equal to the threshold value and took the benefit of threshold voltage with variation of temperature. As we decreases the voltage the value of threshold increases which more under drive the WL at higher temperature. Variation of WL with process corners is shown in Fig.23.



IX. MEMORY WRITE



This is most popular technique to improve write margin of memory cell. In this techniques we improved both write failure like less D.C conditions or improve the strength of pass gate by increasing the D.C voltage and also increase the time slot to improve write failure due to less time slot. By using the strong write driver put the bit-line to ground and then turned off the write driver. After that by using capacitor can put the bit-line below the ground level by using the concept of capacitor i.e. capacitor tries to oppose the sudden changes of voltage across it. So as we sudden change the input voltage (WR) of capacitor from high value to low, capacitor tries to oppose the change of voltage and gives negative bump at output of capacitor which is bit-line or source terminal of pass gate. By using this negative voltage improved the strength of pass gate as compare to pull-up transistor. The strong pass gate makes the fast discharge of node 1 which helps to improve write margin of memory cell.

This technique used to improve the write margin of the memory cell by decreasing the strength of pull-up transistor as compare to the pass gate. Node 1 try to discharge through pass gate for successful of write operation if the strength of

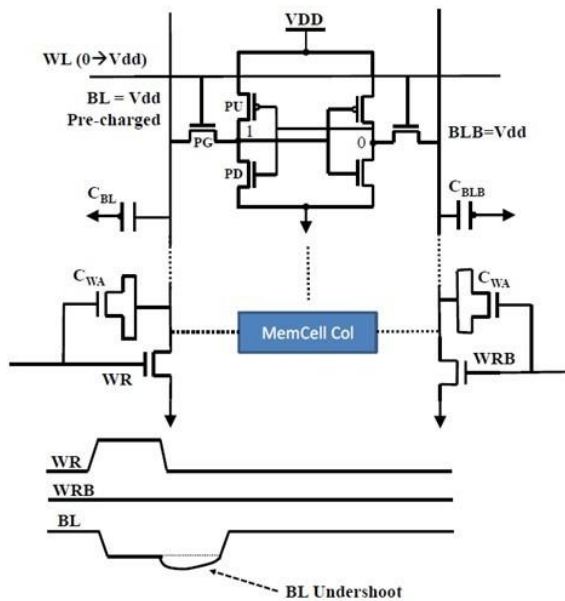


Fig. 25. Negative Bit Line

pull-up transistor is high as compare to pass gate the node 1 never goes to ground which result in write failure. Due to the high strength of pass gate discharging through pass gate is high as compare to the charging of pull-up transistor so node 1 discharge fastly through pass gate and write operation happened successfully.

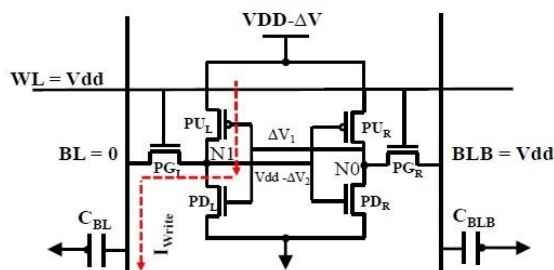


Fig. 26. Lowering Array Voltage

C. Rising of Memory Array Ground Voltage

By increasing the voltage of memory cell rather than putting on ground is another way to improve the write stability of memory cell. In this technique rise the voltage of memory cell by using transistor as a resistor, so the voltage of virtual ground rises and due to high strength of PD transistor, it transfers the voltage of rises virtual ground voltage to node N0 it's mean drain voltage of PU transistor so the strength of PU transistor decreases without compromising the strength of PG. So now the strength of PG is high as compare to the strength of PU transistor and can write the data without write failure. The systematic arrangement of this technique is shown in fig as below.

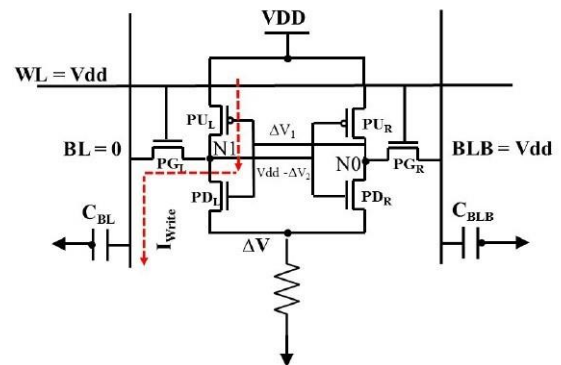


Fig. 27. Rising Ground Voltage

These above techniques show the behavior of write assist circuit which require during the successful operation of write at low voltage. But the other hand due to word-line under-drive or to improve read through read assist circuitry write operation is not happening properly so to avoid the conflict between read and write operation use the assist circuitry simultaneously which improve write as well as read operation of memory cell. The leakage, read, write operation and also cost is improved assist techniques benefits is described in [13]. As voltage is reducing in case of WM may assist techniques is required or when combined operation means read and write operation is going on simultaneously then some assist technique is required [14]. Due to reliability issue at high voltage can be improve by using two phase scheme at low voltage [15]. Another technique is used to improve the write margin at low voltage is capacitive coupling [16]. In this scheme initially BL discharge to ground through write driver and the input of capacitance is kept at pre-charge state. The pre-charge voltage starts to discharge to ground the negative voltage meet at output of capacitance which increase the write current through pass gate. But at high voltage reliability issues occurs due to high negative voltage at source of PG and the voltage increase even high as compare to supply voltage which decrease the life of memory and also can breakdown the thickness of oxide or increases the stress level at the gate input of pass gate. The failure of dielectric oxide is described in[17][18].

X. RELIABILITY ISSUES AND IMPROVED CIRCUIT AT HIGH VOLTAGE OF NEGATIVE BIT-LINE WRITE ASSIST

As technology is shrinking more and more transistor can accumulate on single die but the problem occurs due to static leakage increases as compare to dynamic leakage. So to decrease the static leakage we need to decrease the supply voltage. But due to decreasing the supply voltage stability problem arises in memory and stability also decreases due to conflicting of memory read and write operation. So to overcome this problem and increase the stability of memory cell we require some assist techniques. To improve the write stability generally we used negative bit-line write assist technique which increases the strength of pass gate as discuss above. At the BLR and BLL used strong write driver to put

either BLR or BLL on ground and also used a transistor which works like a capacitor to give negative bump at source of pass gate to increase the strength or for successful write operation. So as we increase the voltage the negative bump given by capacitance also increases which put the gate voltage of PG even higher than the supply voltage. Which is undesirable or increase the stress level of PG or even can break the oxide thickness with such a high electric field [14]. Or decreases the life of a transistor.

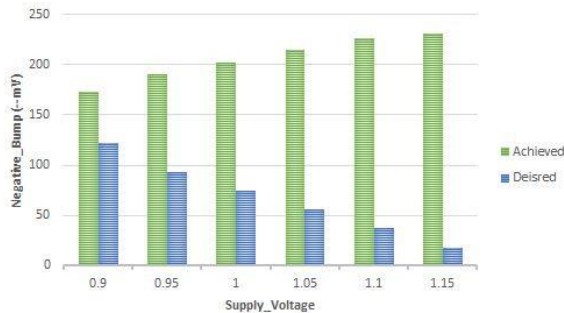


Fig. 28. Negative Bump at low and High Voltage

Also if we increased the voltage more negative bump at the output of capacitance which slow down the speed of write operation at high voltage and also increases the dynamic power. Time dependent oxide [13] also dependent on temperature means when we increase the temperature there is a high chance break the oxide at given time frame as compare to lower temperature on same voltage. Leakage also increases at high voltage due to large negative voltage at BLR or BLL and there may be chance that it can on the unselected memory cell column or increase the leakage due to rise in voltage V_{gs} of PG unselected column.

During write operation at low voltage there is a need of extra assist circuitry but as we increase the voltage no need of assist circuit or depend on the read assist WLUD.

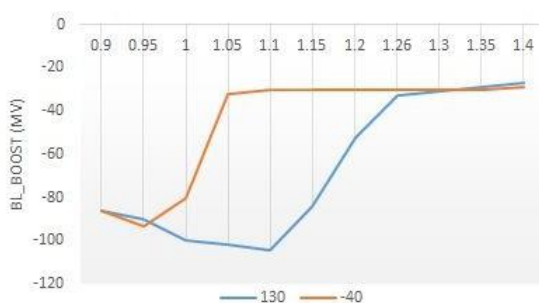


Fig. 29. Improved Negative Bump at High Voltage

XI. CONCLUSION

We designed a low leakage circuit to decrease the leakage current of 6T SRAM memory cell. The leakage current is decreases from 0.25 to 0.22 pico-Amp, as we can see from

the fig.18. And also the stability is improved at SF/125 corner. Read stability is also improved with the help of proposed circuit, the strength of pass gate is decreases as per requirement of process variation and WL under drive voltage value is matched with the requirement value and performance like dynamic power improved by 1.5 percent as compare to previous circuit.

REFERENCES

- [1] A. Kumar et al, "A 6T-SRAM in 28nm FDSOI technology with V_{min} of 0.52V using assisted read and write operation", ICICDT 2015.
- [2] Density 6T SRAM in 40nm CMOS Technology using Adaptive Source Bias"IEEE International Conference on VLSI Design, 2018, pp. 261 -265.
- [3] A. Kumar et al, "Sleep Circuit for SRAM Core with Improved Noise-Margin", IEEE International Conference on Integrated Circuit Design and Technology, 2008, pp. 139-142.
- [4] Nobuaki Kobayashi et al, "A high stability, low supply voltage and low standby power six-transistor CMOS SRAM", The 20th Asia and South Pacific Design Automation Conference, 2015, pp. 10-11.
- [5] A. Kumar et al, "A Temperature Compensated Read Assist for Low V_{min} and High Performance High Density 6T SRAM in FinFET Technology", VLSID 2018.
- [6] Y Takeyama et al, "A low leakage SRAM macro with replica cell biasing scheme", IEEE Journal of Solid-State Circuits (Volume:41 , Issue: 4 2006, p.p. 815-822.
- [7] Cyrille Dray et al, "A 40nm low power SRAM retention circuit with PVT-aware self-refreshing virtual VDD regulation", IEEE International Memory Workshop, 2010, p.p. 1-4.
- [8] Nobuaki Kobayashi et al, "A high stability, low supply voltage and low standby power six-transistor CMOS SRAM", The 20th Asia and South Pacific Design Automation Conference, 2015, pp. 10-11.
- [9] Fatih Hamzaoglu et al, "A 3.8 GHz 153 Mb SRAM Design With Dynamic Stability Enhancement and Leakage Reduction in 45 nm High-k Metal Gate CMOS Technology" IEEE Journal of Solid-State Circuits, Volume 44, Issue 1, 2009, p.p. 148-154.
- [10] Peter Kuoyuan Hsu et al., "A SRAM cell array with adaptive leakage reduction scheme for data retention in 28nm high-k metal-gate CMOS", Symposium on VLSI Circuits (VLSIC) 2012, pp 62-63.
- [11] Yoshisato Yokoyama, et al., "40nm Ultra-low leakage SRAM at 170 deg.C operation for embedded flash MCU", Fifteenth International Symposium on Quality Electronic Design, 2014, pp. 24-31.
- [12] L. Chang, D. Fried and J. Hergenrother, "Stable SRAM cell design for the 32 nm node and beyond", Symposium on VLSI Technology, pp.128- 129, 2005
- [13] Vikas Chandra, Cezary Pietrzyk, and Robert Aitken, "On the Efficacy of Write-Assist Techniques in Low Voltage Nanoscale SRAMs", Design, Automation and Test in Europe Conference and Exhibition (DATE), 2010, pp. 345-350.
- [14] Farah B. Yahya, Harsh N. Patel, Vikas Chandra, and Benton H. Calhoun, "Combined SRAM Read/Write Assist Techniques for Near/Sub-Threshold Voltage Operation", Asia Symposium on Quality Electronic Design, 2014.
- [15] M Sultan M Siddiqui, Shailendra Sharad, Yogendra Sharma and Amit Khanuja, "Two Phase Write Scheme to Improve Low Voltage Write-Ability in Medium-Density SRAMs", International Conference on VLSI Design (VLSID), 2015, pp. 176-180.
- [16] Jaydeep Kulkarni, Bibiche Geuskens, Tanay Karnik, Muhammad Khel-lah, James Tschanz, and Vivek De, "Capacitive-Coupling Wordline Boosting with SelfInduced VCC Collapse for Write V_{min} Reduction in 22-nm 8T SRAM", International Solid State Circuit Conference (ISSCC), 2012, pp. 234-235.
- [17] D. P. Wang, H. J. Liao, H. Yamauchi, Y. H. Chen, Y. L. Lin, S. H. Lin, et al., "A 45 nm dual-port SRAM with write and read capability enhancement at low voltage," in Proc. IEEE Int. SOC Conf., Sep. 2007, pp. 211-214.
- [18] A. Haggag, M. Moosa, N. Liu, D. Burnett, G. Abeln, M. Kuffler, et al., "Realistic projections of product fails from NBTI and TDDB," in Proc. 44th Annu. IEEE Int. Rel. Phys. Symp., Mar. 2006, pp. 541-544.