

A Study of Circuit Level Leakage Reduction Techniques in Cache Memories

Urvashi Chaudhari*, Rachna Jani**

*(Department of Electronics & Communication, CHARUSAT University, Changa)

** (Associate Professor Department of Electronics & Communication, CHARUSAT University, Changa)

ABSTRACT

The Performance of microprocessor can be improved by increasing the capacity of on-chip caches. On-chip caches consume noticeable fraction of total power consumption of microprocessors. The performance gained can be achieved by reducing energy consumption due to leakage current in cache memories. The technique for power reduction in cache is divided in mainly two parts Circuit level and architectural level technique. In this paper a circuit level techniques like gated-V_{DD}, gated-Ground, Drowsy caches, Asymmetric SRAM cell for reducing leakage current in cache memory are discussed.

Keywords - Asymmetric cell, Cache memory, Drowsy caches, Gated-ground, Gated-V_{DD}, leakage current.

I. INTRODUCTION

A Microprocessor devotes a large fraction of chip area for memory structure. Due to large size of on-chip caches reduction of leakage current even in single cell of cache can reduce a large fraction of the total power in the microprocessor. The main memory compared to cache memory is slow and not able to maintain the speed with processor. Cache memory is placed between the main memory and the processor. SRAMs are used as a cache memory because they are faster and not required periodic refresh compared to DRAM. Now leakage current becomes a major contributor for power dissipation in CMOS circuit in deep submicron technology. So, reduction in leakage current becomes main task for designers. It is also important to estimate leakage current in CMOS at nanotechnology. An efficient technique for estimating leakage current has been proposed in [1]

Many circuit and architecture level techniques have been proposed for leakage reduction. The architectural level technique can reduce power significantly but it produce negative impact on performance. So that in this paper some of circuit level techniques are discussed which reduces leakage and less impacts on performance. There are many circuit level techniques for leakage current reduction. One of them is multi-threshold CMOS (MTCMOS) technique. It increases operating speed

by low- V_{TH} MOSFET and leakage current by high- V_{TH} MOSFET [2]. This technique use high- V_{TH} pMOSFET and nMOSFET as a Switch for disconnecting power supply in standby mode and thereby reduce leakage current. Disadvantage of this circuits are fabrication of MOS with different threshold-voltage, increase overall circuit area and it adds extra parasitic capacitance and delay. Another technique is Dual V_{TH} [3]. It uses low V_{TH} for critical path and high V_{TH} for rest of the circuit. Due to high V_{TH} , it increases the access time of the memory cell. Dynamic sleep transistor also used as leakage reduction technique [4]. In that sleep transistor is use to isolate the circuit from the supply.

The rest of the paper organized as follows: Section II describes Gated- V_{DD} techniques. Section III describes Data retention gated-ground cache. Section IV describes Drowsy cache. Section V describes Asymmetric SRAM cell. Section VI describes Conclusion.

II. GATED-V_{DD}: GATING THE SUPPLY VOLTAGE

To prevent leakage power dissipation in DRI i-cache [5] a circuit level technique, gated- V_{DD} is used, by gating the supply voltage from the SRAM cell of the cache unused portion [6].

A. SRAM Cell with gated- V_{DD} :

In this techniques an extra transistor in the supply voltage or the ground of the cache's SRAM cell is added as shown in fig.1, it turn on in the used section and turn off in unused section, so the cell's supply voltage is gated. The main reason behind the reduction in leakage current is stacking effect of self reverse-biasing series-connected transistors.

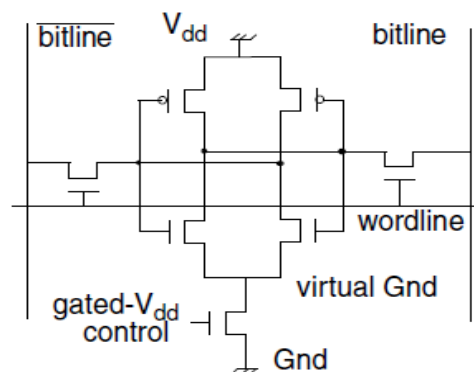


Fig.1. SRAM cell with an nMOS Gated-V_{DD} [6].

This technique maintains both lower supply and threshold voltage while reducing leakage current. Gated-Vdd transistor must be large enough to sink the current through SRAM cell, but a large transistor may reduce the stacking effect and also increase the area overhead. So here a trade-off between area overhead and leakage reduction.

B. Trade- offs between nMOS and pMOS Gated Vdd transistor:

By Using nMOS transistor as Gated-Vdd, it reduce standby leakage current through stacking effect of three series connected nMOS transistor between bitlines and ground. Alternatively, using pMOS transistor it reduces required transistor width and thereby reduce area overhead, and also it not provide the isolation between the bitlines and the ground as nMOS, reducing energy saving. nMOS gated-Vdd impact on cell performance while pMOS not significantly impact on the cell performance. There is a fundamental trade-off between reduction in leakage, transistor switching speed and area overhead of gated-Vdd transistor.

C. Impact of lowering threshold voltage:

In table 1 from first three rows it is concluded that decreasing threshold voltage of SRAM cell increases active leakage energy and standby leakage energy. From last three rows, if threshold voltage of gated-Vdd decrease than there is further reduction in the standby leakage energy.

Table 1. Impacts of changing SRAM and gated-Vdd threshold Voltages [6]

SRAM Cell V_t (V)	Gated-V _{dd} V_t (V)	Active Leakage Energy (nJ)	Standby Leakage Energy (nJ)
0.40	0.40	12	10
0.30	0.40	143	49
0.20	0.40	1700	50
0.40	0.20	12	11
0.30	0.20	143	76
0.20	0.20	1700	165

D. Impact of widening gated-Vdd transistor:

Table 2. Widening the gated-Vdd transistor [6]

Area Increase (%) of NMOS Gated-V _{dd}	Relative Read Time	Active Leakage Energy (nJ)	Standby Leakage Energy (nJ)
2	1.00	1700	166
4	0.90	1710	245
8	0.85	1720	371

From table 2 it is shown that by increasing the size of the gated-Vdd transistor read time of the circuit decreases but active and standby leakage energy increases. For a processor this technique without

data retention capability reduces the leakage power by 97% while increasing read time by 8% using 0.2v and 0.4v for low and high V_{TH} respectively [7].

III. Data Retention Gated-Ground Cache (DRG-Cache):

In this technique an extra nMOS transistor connected between ground and virtual ground node of SRAM cell, which is called gated-ground. In this technique, to reduce power, the unused section of the memory is put in to the low leakage mode. Gated-ground transistor enables a DRG-cache to turn off the supply voltage and eliminate leakage energy virtually in unused section of the memory [8].

A. DRG-Cache Technique with SRAM Cell

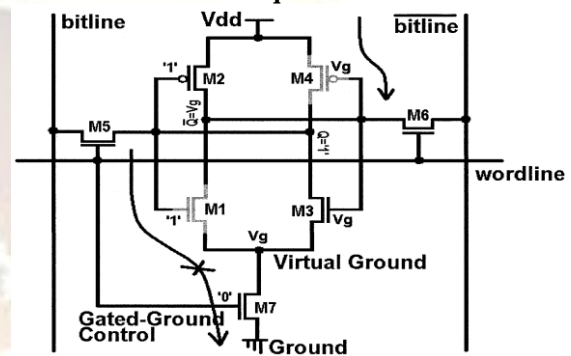


Fig.2. Circuit of gated-ground transistor with SRAM cell [8].

In fig. 2 light colour transistor are on and dark colour transistor are off. The extra transistor turns on in the used section and turn off in unused section of the memory thereby gating the supply voltage. As in the gated-Vdd technique, this technique reduces leakage due to series connection of two off transistor. This effect is due the stacking effect. Here this technique retains the data which is not possible in the case of gated-Vdd. A careful design of transistor is necessary because size of the gated-ground is important in data retention and stability.

B. Energy Performance Trade-off:

Table 3. Energy performance trade-off [8]

Gsize	V _{tn0} (V)	Normalized Conv. Cache leakage	Normalized DRG Cache leakage	Relative Read Time
2.0	0.45	1.0	0.68	1.028
1.0	0.45	1.0	0.54	1.046
1.0	0.35	9.0	5.2	1.044
1.0	0.25	80.8	47.8	1.044

First two rows of table3 shows that increasing the width of the gated-ground transistor improve the read time, data retention capability and stability of cell, but decrease the energy saving and increases the area. Last two row shows that threshold voltage increases the leakage energy. This

technique for a processor with V_{TH} 250mv reduce leakage power by 40% while increase read time by 4.4% compared to conventional Cache [7].

IV. DROWSY CACHES:

In Caches, for fix period of time the activity is centred at some cache lines. So, putting rest of cache lines in low power mode can reduce the leakage significantly. This low power mode of cache line is called Drowsy Caches [9][10]. In stand of turning off cache line putting it in to a low power drowsy mode can reduce leakage significantly. In drowsy caches the chance of putting wrong line into drowsy mode is less, for that different policies have been proposed in [9]. When caches are in drowsy mode the data in it are preserved. Drowsy caches can be implemented by adaptive body-biasing with multi-threshold CMOS (ABB-CMOS), dynamic voltage scaling (DVS). Gated-Vdd.

A. Drowsy memory circuit:

As shown in fig.3 SRAM cell is connected to voltage-scaling controller. This controller consist of two pMOS pass transistor, one with high threshold voltage while other with low threshold voltage. One pMOS supplies normal supply voltage and other low for drowsy cache lines. Each pass transistor of SRAM cell is of high V_{th} to prevent the leakage current from the normal supply to the low supply through the two pMOS pass gate transistor. For each cache line a separate voltage controller is needed.

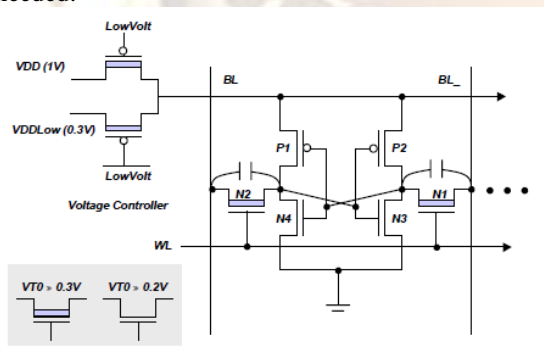


Fig.3. Schematic of Drowsy memory circuit [9].

A possible disadvantage of this circuit is that increased susceptibility to noise. In a 0.07um CMOS process, drowsy caches will be able to reduce the total energy consumption by 50% -75%, and cache lines can be maintained in drowsy mode without affecting performance by more than 1% [9].

V. Asymmetric SRAM Cell:

As this technique used in cache it is refer to as asymmetric-cell caches (ACCs). Comparing to conventional cache ACCs reduce leakage power even when there are few parts of the cache that are left unused [11]. Traditional SRAM cell transistors are symmetrical with identical leakage and threshold

voltage, while asymmetric SRAM cells have low leakage and less impact on performance. In this technique when cell is storing '0', selected transistors are "weakened" to reduce leakage. A weakening can be possible by using higher threshold voltage and also by proper sizing of transistor.

In conventional SRAM of symmetrical transistor to reduce leakage current one method can be used that is making all transistors of high V_{th} , but it degrades the performance. This drawback can be overcome by using asymmetric SRAM cell. It works on following principle: selecting a preferred stored value and weaken only those transistors necessary to reduce leakage by increasing the threshold voltage when this value is stored.

A. Working of Asymmetric SRAM Cell:

In cell most of the leakage is dissipated by transistors that are off and have a voltage differential across their drain and source. This state of transistor can be finding by the value stored in it. When a cell storing a '0' value, as shown in fig.4, the leaky transistors will be P1, N4 and N2. If cell storing '1' value then leakage transistor would be P2, N1 and N3.

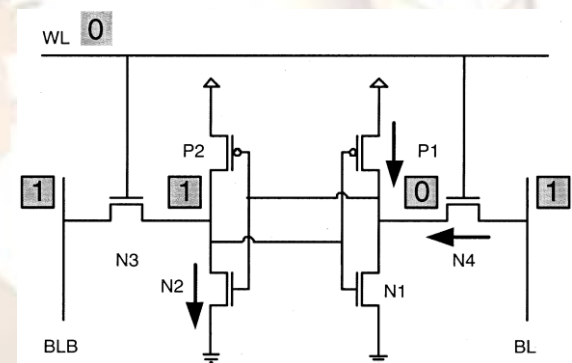


Fig.4. SRAM Cell with storing a '0' value [11].

To reduce leakage in cell when it storing a '0' value, replace leaky transistor by high V_{th} . The resulted circuit is shown in fig. 5, which is called basic asymmetric (BA) SRAM cell. This circuit has the same leakage as conventional SRAM cell when storing '1', but it reduces leakage by 70 times when storing '0', due to longer discharge time.

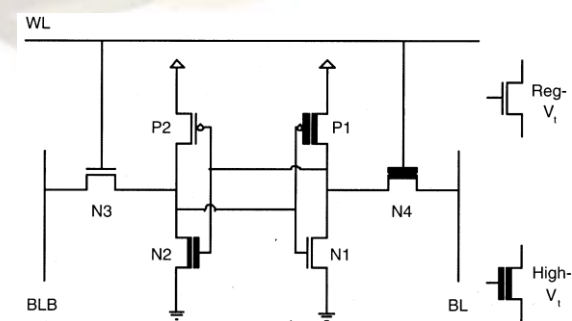


Fig.5. Basic asymmetric SRAM cell [11].

To improve in leakage reduction and speed other modified circuit of basic asymmetric SRAM cell is proposed in [11]. In that they have proposed two best circuit designs for improvement in leakage reduction and speed from different asymmetric SRAM cell family. For leakage reduction improvement, leakage enhancement (LE) cell and for speed improvement, speed enhancement (SE) cell. Leakage in LE cell for storing '0' and '1' is 1% and 14% respectively, and that for SE cell is 14% and 50% respectively. Expected leakage and delay in LE cell are 5% and 2% respectively, and that in SE cell are 25% and 0% respectively [11].

VI. CONCLUSION

Table 4. A comparison of all techniques:

Techniques		Leakage reduction	Impact on performance
Gated-Vdd		97% with no data retention	8%
Gated-ground		40%	4.4%
Drowsy cache		50 %-70%	Not more than 1%
Asymmetric SRAM Cell	LE Cell	95 %	2%
	SE Cell	75%	0%

The comparison of all technique is shown in Table4.From this comparison it can be concluded that asymmetric cell and drowsy cache are better techniques for leakage reduction in cache. Though asymmetric cell have high leakage reduction and better performance, it has some drawbacks. Asymmetric cell do not consider gate leakage component so that their total leakage saving will be less. However asymmetric cells are expected to considerably decrease the static power dissipation at high operating temperature. Furthermore in successive technologies, the stability of asymmetric cells may decrease.

REFERENCES

[1] Alodeep Sanyal, Member, IEEE, Ashesh Rastogi, Wei chen and Sandip Kundu, fellow, IEEE, "An efficient technique for leakage current estimation in nanoscaled CMOS Circuits incorporating Self-Loading Effects" *IEEE TRANSACTION ON COMPUTERS*, VOL 50,NO 7,JULY 2010.

[2] S. Mouth, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, "1-V power supply high-speed digital circuit technology with multithreshold- voltage CMOS," *IEEE J. Solid-State Circuits*, Vol 30,pp. 847~854, Aug.1995.

[3] J.Kao, and A. Chandrakasan, " Dual-threshold voltage techniques for low power digital circuits," *IEEE J. Solid-state Circuits*,vol 35, pp. 1009~1018, Jul. 2002

[4] K.Zhang, U. Bhattacharya, Z. Chen, Faith, D. Murray, Member IEEE, " SRAM design on 65nm CMOS Technology with Dynamic sleep Transistor for leakage reduction", *IEEE J. Of solid-state circuits*, vol.40, no.4, April 2005.

[5] S.H. Yang, M.D. Powell, B. Falsafi, K.Roy, T.N. Vijaykumar, " *Dynamically resizable instruction cache: An energy efficient and high-performance deep-submicron instruction cache.*" Technical report ECE-007, School of Electrical and Computer Engineering, Purdue University,2000.

[6] M.powell,S.Yang, B Falsafi, k.Roy, T Vijaykumar, "Gated-VDD: A circuit technique to reduce leakage in deep-submicron cache memories," *porc. Of IEEE/ACM Intl symp. On lower power electronics & Design*, pp 90~95,2000.

[7] N.Sung Kim, K. Flautner, D. Blaauw, and T. Mudge, "Circuit and Microarchitectural technique for Reducing Cache Leakage Power", *IEEE Trans. On Very Large Scale Integration Systems*, Vol.12, No.2, February 2004.

[8] Amit Agarwal, Hai Li, Student Member, IEEE, and Kaushik Roy, Fellow, IEEE, "A Single-Vt Low-Leakage Gated-Ground Cache for Deep Submicron", *IEEE J. Of Solid-State Circuits*, Vlo.38, no.2, February 2003.

[9] K. Flautner, Nam sung kim, S. Martin, D. Blaauw and T. Mudge, "Drowsy Caches: simple techniques for reducing leakage power," *proc. Of IEEE/ACM Intl.Symp. on computer Architecture*, PP. 148~157, 2002.

[10] Nam sung kim, K. Flautner, D. Blaauw and T. Mudge, "Drowsy instruction cache-Leakage power reduction using dynamic voltage scaling and cache sub-bank prediction." *proc. Of IEEE/ACM Intl.Symp. on Microarchitecture (MICRO-35)*, pp.219~ 230, Nov 2002.

[11] Nvid Aziz, student member, IEEE, Farid N. Najm, fellow, IEEE, and A.Moshovos, Associate member,IEEE, " Low-Leakage Asymmetric-Cell SRAM", *IEEE Trans. On Very Large Scale Integration System*, vol.11, no.4. August 2003.