LOW-POWER HIGH-SPEED CIRCUIT DESIGN FOR VLSI MEMORY SYSTEMS

Naresh Reddy B[1], Dr. Yash Pal Singh[2]

Department of Electronics and Communication Engineering

[1,2]OPJS University, Churu (Rajasthan), India

Abstract

In present day PC systems, progressive memory engineering is broadly utilized. There are different memory components inside one system, for example, enrol records, cache memories, and basic memories. Rapid memory system design has been and will have been a standout amongst the most imperative outline issues. As systems go toward higher execution, threshold of these memories gets bigger. In microchips, for instance, on-chip cache sizes are developing with every era to connect the expanding dissimilarity in the velocities of the processor and the principle memory. Power dispersal has likewise turned into an essential thought both because of the expanded combination and working rates, and additionally because of the hazardous development of battery worked apparatuses.

1. INTRODUCTION

Process variety, regardless of presenting challenges in steadiness and energy of VLSI recollections, gives another chance to spillage control diminishment: measurable reenactment demonstrates that the same VLSI cell spills diversely while putting away 0 and 1; this distinction is as high as 57% at 60mv variety of limit voltage (Vth). Hence, spillage can be decreased if values with less spillage can be put away in the cells. We demonstrate relevance of this proposition by introducing a first procedure for diminishing guideline cache spillage: we reorder directions inside essential squares to coordinate the guidelines with the less-broken condition of their relating cache cells. Trial comes about appear to 12.51%, averaging 10.73%, spillage energy decrease at 60mv variety in Vth and that this sparing increments with innovation scaling. Since intra-essential piece rescheduling does not influence direction cache hit proportion, this diminishment is given just a unimportant punishment, in uncommon cases, in the information cache.

In plan of complex number juggling rationale circuits, ground bob commotion, standby spillage current and spillage control are essential and testing issues in nanometer down scaling. In this paper, a low power, low intricate and lessened ground bob commotion full adder configuration in view of pass transistor rationale (PTL) is proposed. Essentially adder is imperative piece of complex number juggling rationale circuit in number-crunching operation like expansion,

augmentation and computational model. Adder circuit is generally utilized as a part of numerous advanced circuits for number-crunching operation as well as address era in processor and microcontroller recollections which are utilized in expansive scale framework at higher speed .In this paper, we have proposed an altered 10T full adder in view of PTL utilizing multi-edge CMOS method.

Here we use forward body one-sided multimode (MTCMOS) system to assess standby spillage current, power and ground bob clamor. All the reenactment in this paper has been done utilizing rhythm virtuoso at 45 nm innovation at different voltage and temperatures. The forward body one-sided (FBB) multimode MTCMOS procedure has been executed on regular 10T full adder circuit with 45 nm innovation parameters for reproductions. By utilizing this strategy the standby spillage current decrease can be enhanced by 90 % and spillage energy to 30 % when contrasted with base 10T full adder. Ground bob clamor can be lessened to 55 % when contrasted with the traditional adder. Inside bite the dust procedure variety is expanding in nanometer-scale process advancements.

2. INPUT/OUTPUT CIRCUIT

Chip-to-chip correspondence through I/O is both speed bottleneck of the system and an overwhelming wellspring of power since a capacitive heap of outside hubs is bigger roughly by three requests than that of interior hubs. Along these lines, fast and low-power interface circuits have pulled in much consideration.Figure1 shows power pattern of MPUs and DSPs arrayed at International Solid-State Circuits Conference (ISSCC). Power utilization of top of the line processors is currently more than 100W. Today, low power is an issue for convenient applications as well as a stringent interest for fast applications.

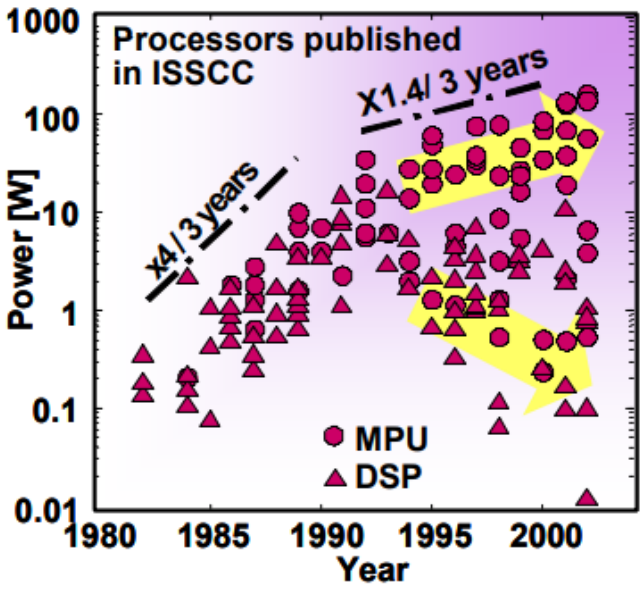

Figure 1: Power Trend of MPUs and DSPs Presented ISSCC

For the most part, a system has two sorts of power utilization. One is dispersed inside a chip by rationale circuits, timing circuits, and on-chip memories. The other is scattered by I/O circuits when at least two chips inside a system transmit and get data. Figure 2 shows power appropriation of four distinctive CMOS LSIs. As is comprehended from the figure, power dissemination turns out to be entirely unexpected relying upon applications, however, it ought to be noticed that the total of memory and I/O power regularly possess a huge bit of the total power. Considering the way that fasts I/Os between a processor and primary memories is the most eager for power I/O, one might say that low-power and rapid memory systems are required practically in all systems and applications.

Rambus interface is one of the illustrations which are broadly utilized between a microchip and DRAM based primary memories. Current interface strategies, in any case, are not perfect answers for the future to come.

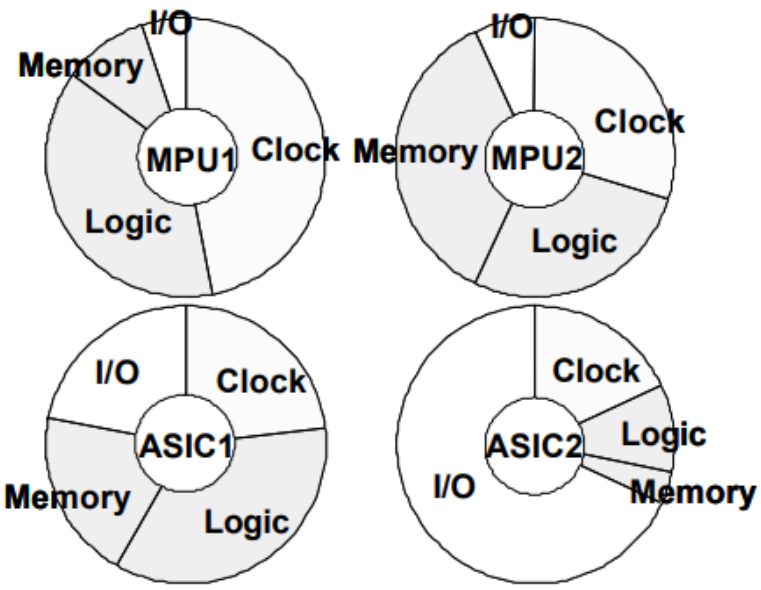

Figure 2: Power Distribution in CMOS LSIs

Table 1 shows several important parameters of future high-performance LSIs predicted by SIA [1]. In the year 2016, when the mainstream technology is shrunk down to 22nm gate length, the maximum on-chip clock frequency is 29GHz, microprocessor power is 288W with 0.5V supply voltage, and number of I/O pins is 7100. These goals cannot be reached by simply taking advantage of scaling scenario [2], since scaling of I/O related circuits is slower than that of core logic circuits. The main reasons lie in difficulties of packaging and also in special design considerations regarding I/Os such as ESD protection.

	Unit	Year 2002	Year 2016	Factor
Design rule	nm	130	22	0.2
Transistor density	/cm^2	112M	2.8G	25
Chip size	mm^2	800	572	0.7
Tr. count per chip (µP)		900M	16G	20
Local clock on a chip	Hz	2.3G	29G	13
Power	W	140	288	2.1
Supply voltage	V	1.2	0.6	0.5
# of package pins in ASICs		1870	7100	3.8

Table 1: Key Parameters of Future LSIs Predicted by SIA

Table 1 shows four distinctive existing interface plans, rapid serial connections, 3D

coordination, installed DRAM (eDRAM) and miniaturized scale knocks (µ-knocks). These technologies are looked at from different stance of view in Table 1. Despite the fact that fast connections are generally utilized, it has an issue in the adaptability of power utilization and is additionally hard to have 7100 I/O cushions because of moderate scaling of I/Os. Utilization of little swing and impedance-coordinated interface requires very simple plan methods which Time Turnaround time (TAT) and cost. 3D reconciliation can develop perfect interface by utilizing MEMS technology.

Realistic Synthesizer chip manufactured in 0.18µm CMOS technology, for instance, has 32 Mb on-chips DRAM attached to rationale circuits with 2560 piece wide transport, and accomplished 48GB/s correspondence transfer speed at 150MHz. Since correspondence between a processor and a DRAM is currently finished with on-chip transports, more extensive data transfer capacity and higher operation frequency are accomplished with reduced power contrasted and off-chip interchanges like serial connections.

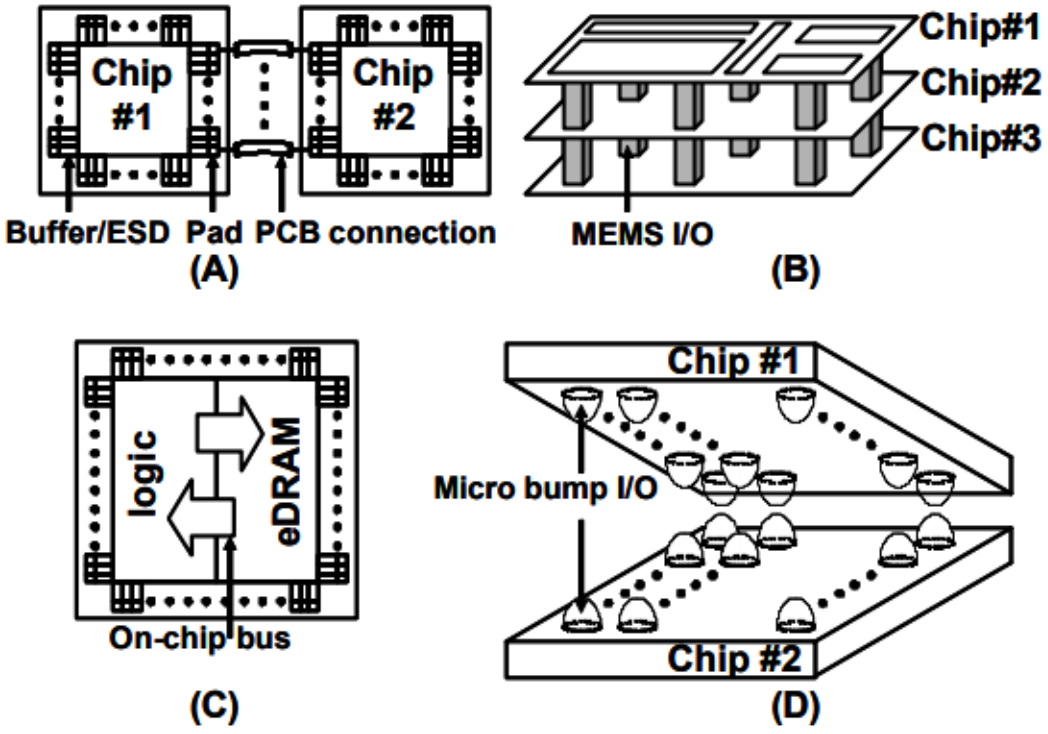

Figure 3: Various Existing I/O Schemes. (A) High-Speed Serial Links on a Printed Circuit Board. (B) 3D Integration. (C) Embedded DRAM. (D) Micro Bump

By using the concept of SOC, various components can be implemented on one chip to allow the chip to have high functionality and performance. Figure 3 shows mask count increase for integrating various circuit blocks on a chip. Though SOCs have attracted much attention for

future system LSIs, it is understood from the figure that integrating everything on the same die is not a realistic solution for all applications due to high mask count, design and process complexity, and therefore cost. Increased chip area also lowers production yield and increases cost. Building blocks like analogy circuits and flash memories are difficult to be integrated on the same die as logic due to noise and process problems. Possibility of success in first silicon becomes low and test process becomes very complex. Therefore, SOCs also have only threshold applications whose market is large enough that high production cost can be surely got back.

	(A) High-speed serial links	(B) 3D integration	(C) eDRAM	(D) μ-bumps
Detachability	-	-	-	-
Bandwidth	-	+	++	+
Chip stacking	N/A	8	N/A	2
Testability	++	-	++	-
Durability	-	+	+	+
Cost	++	--	-	+
Usability	-	N/A	N/A	N/A
Area	-	++	+	+
ESD	-	-	-	-
Power	--	++	+	+
TAT	-	---	--	+

Table 2: Comparison of Four Interface Schemes

System in-package (SIP) is an option answer for SOCs that can assuage cost issue. Figure 4 shows applied the perspective of SOCs and SIPs. Utilizing SIP, LSI creators require not stoop everything without anyone else, but rather they can utilize different IPs from other LSI organizations. Furthermore, by utilizing IPs, different ASICs can be planned inside a brief span to address market's issues. Multi-chip module (MCM) technology has turned out to be mainstream with the reason expressed previously. μ-knocks [3] are one of the SIP arrangements. I/O cushion is shaped to the highest point of an LSI chip with a thickness of 30μm pitch.A microchip and a DRAM are put up close and personal as appeared in Figure 4, and they can have vast I/O stick check working at high frequency since electrical length for signals is tremendously lessened contrasted and that of PCB associations.

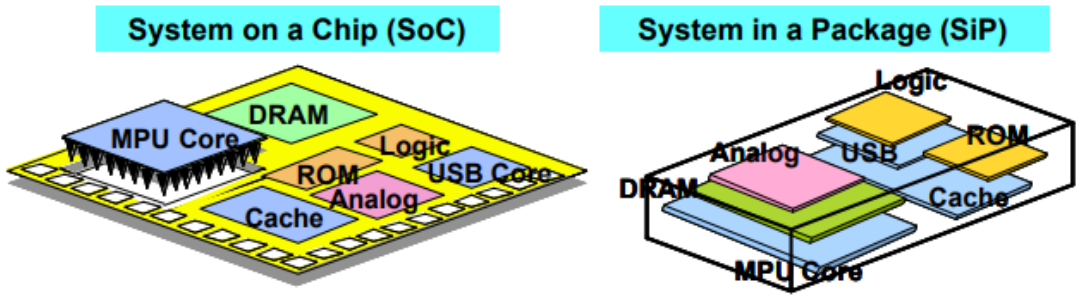

Figure 4: System-On -Chip vs. System-In-Package

That is need of Electro-Static Discharge (ESD) security. Since metal I/O cushion is uncovered on the highest point of the chip for having physical contacts between chips, ESD insurance circuits ought to be introduced for every I/O cushion. ESD structure, for the most part, includes pico-farad of capacitive load on the cushion, which increase postponement, power, and region and thwarts the scaling of the cushions further. From the above exchanges, it is obviously seen new plan methods for rapid and low power I/O is emphatically sought. In the theory, new I/O plot, Wireless Super interfaces (WSC) conspire is depicted for future LSIs. Power devoured by on-chip memory circuits themselves is additionally a basic issue in future LSIs.

3. LOW VOLTAGE OPERATION

Here, theoretical minimum value of supply voltage for a general VLSI cell is given, and it is shown that a CMOS VLSI cell is the best solution for future low-voltage operations. Figure 5 shows a general VLSI cell circuits. Let us assume that "1" is written on the node N2 from BL. Just after a write cycle, data "1" inside the cell is less than VDD due to VTH drop of the access transistor. Then, let us assume that the same cell is accessed in the next read cycle. In order that enough read current IREAD flows, driver NMOS M5 should be turned on. This condition can be expressed as follows.

$$V_{GM5} > V_{TH} \quad (1)$$

V_{GM5} is the gate voltage of M5 and V_{TH} is the threshold voltage of M5. If the load devices are implemented with high-resistive load or T_{FT}, V_{GM5} is expressed as follows.

$$V_{GM5} = V_{DD} - V_{TH}' \quad (2)$$

VTH ' in (2) denotes the threshold voltage of the access transistor including body effects. It takes more than 1000 cycles for the load device to charge the node N2 up to VDD due to its

poor current drivability. This is because the node N2 is charged up to VDD soon after word line goes to low in a write due to the strong current drivability of the load PMOS.

- **Circuits**

Address signals are divided into two categories, row address and column address. Figure 5 shows 2bit row and column decoder respectively. Row decoders and column decoders almost have the same structure except that the output of column decoders should be complementary as inputs to transmission gates for allowing full-swing connection between bit lines and data lines.

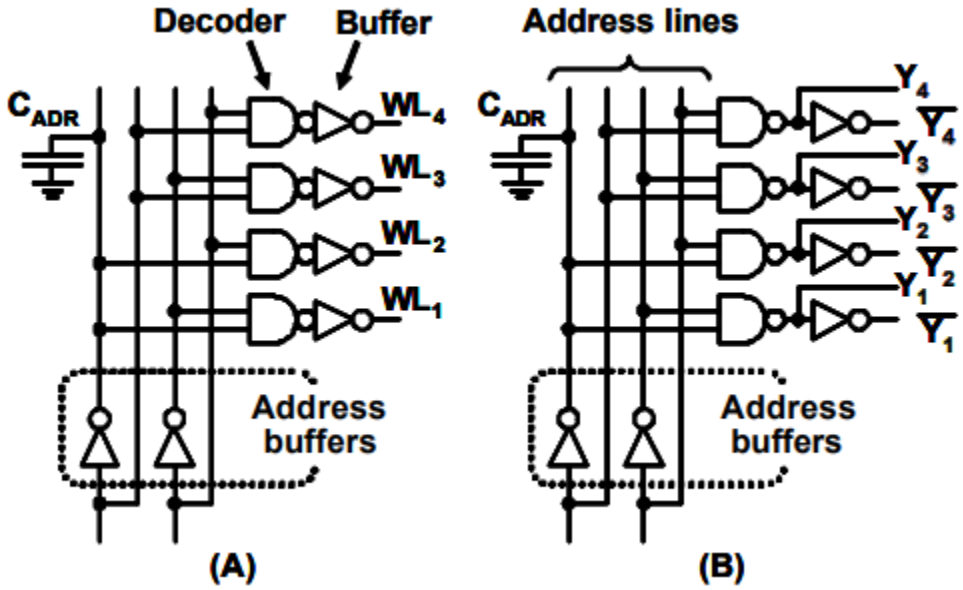

Figure 5: Decoders (2bit) (A) Row Decoder (B) Column Decoder

When number of address bits is N, the number of decoders is 2N. Usually N is as large as 10, and capacitive load of an address line, CADR is high due to the gate input capacitance of many decoders and the line capacitance associated with the long wires. Some alternative decoder circuit styles can be found in [4], which utilize pseudo NMOS and pass transistor techniques to reduce decoder size and capacitance of address lines. When integration size is large, the number of cells connected to a word line is as large as 256, which makes a word line very long and highly capacitive. In addition, resistance of a word line made of poly silicon is high and the associated RC delay may be long. RC delay of a word line can be lowered to half by placing a decoder at the centre of the word line, since both resistance and capacitance of one branch become half.

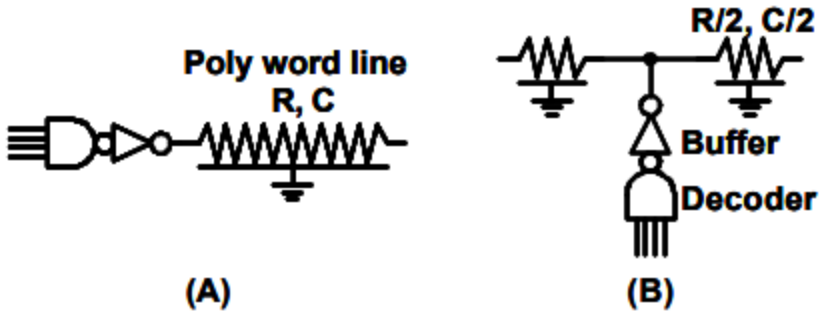

Figure 6: Decoder Placement (A) Driving WL from the Left end (B) Driving WL at the Centre

- **Decoding architecture**

With the increased memory size, the number of cells connected to one word line increases. Consequently the capacitive load and delay become large. In addition, when a word line is high, all cells connected to the word line draw current from bit lines and consume large power. In order to solve this problem, hierarchical structures are introduced which are known as divided word-line (DWL) structure [5] and hierarchical word decoding (HWD). Except the difference in the number of hierarchy level, basically they have the same concept.

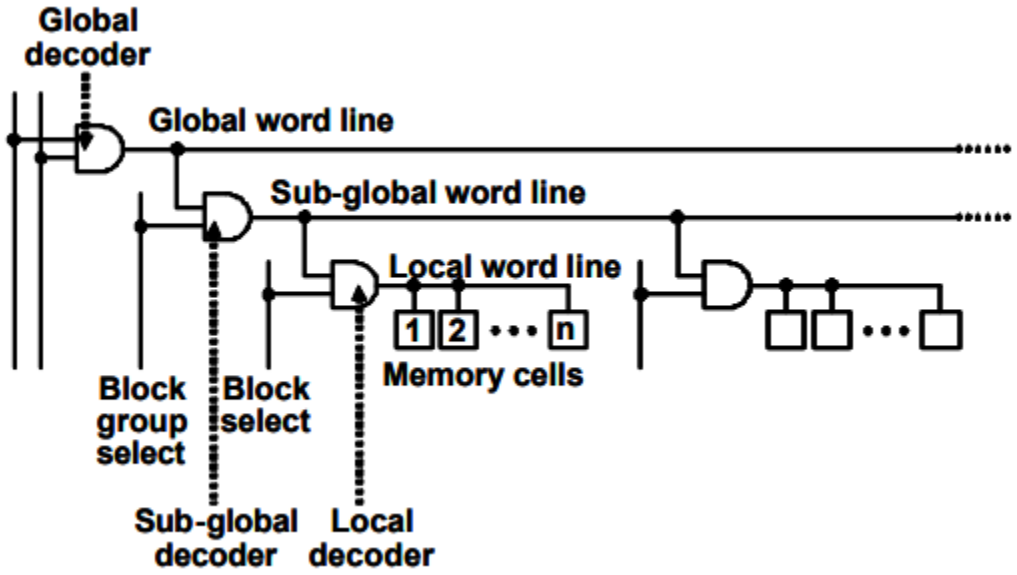

Figure 7: Hierarchical Decoding Structure for Large-Capacity SRAMS

The concept is illustrated in Figure7. One global word line is divided into a group of sub-global word lines, and one sub-global word line is also divided into a group of local word lines. In each

hierarchy, only one word line is activated. Though the total number of decoders is increased, length of each word line is reduced, number of activated features is also reduced, and thus delay and power is improved. Another important design technique is pre-decoding. Address signal lines are distributed to each decoder input. When one of the four word lines is selected, two of the four address signal lines consume active power. When prerecording scheme is used, decoding is done prior to distributing signal lines.

- **Leakage Mechanism**

When one memory cell includes single defect which causes abnormal leakage current, there are two methods to eliminate its path:

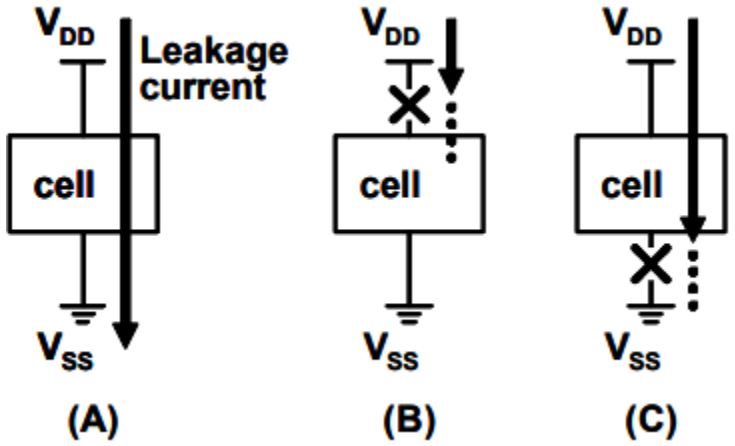

Figure 8: How to Eliminate Abnormal Leakage Current Path

One is to seclude the cell from VSS side, and the other is to confine the cell from VDD side. In VLSI, notwithstanding, the previous approach is hard to acknowledge since the VSS lines for a memory cell cluster are normally organized as a work and are difficult to be cut off specifically. Along these lines, the last way is sought after. In a standby mode, word lines are set to VSS, while bit lines are pre-charged to VDD. At that point, there are just three sorts of ways that pass on VDD to a memory cell in a standby mode.

4. CONCLUSION

Peripheral circuits include sense amplifiers, write buffers, and other timing generators. Describing details of all of these features is beyond the range of this section, and only some basics are given. Sense amplifier is one of the most important circuits which determine both speed and power. There are a number of circuit topologies for sense amplifiers, but they can be divided into two categories, voltage sensing type or current sensing type. Among voltage

sensing type, there are also two kinds of amplifiers, current-mirror type and latch type.VLSI are widely used as key components such as main memories, caches, and buffer memories in high-speed portable systems, where low standby current is strongly required to extend the lifetime of batteries. Power leakage of normal SRAMs in a standby mode is normally negligibly as small as 1μA due to the highly resistive nature of MOSFETs in off-state.

There are numerous routes in which this strategy can be executed, however, the essential thought is to close down the power supply so the site without moving units don't expend any power. This should be possible utilizing some high threshold transistors called rest transistors [6]. In the event that the threshold voltages of rest transistors are changed at runtime, the triple-well technology is required. Plausibility is to utilize different threshold voltage CMOS (MTCMOS) [7]. In MTCMOS, a high threshold gadget is embedded in the arrangement with low threshold transistors making a rest transistor. This makes virtual supply and ground rails whose voltage levels are near the genuine supply and ground lines as a result of the little on-resistance of the embedded high-transistors. Practically speaking, just a single virtual rail (generally the virtual ground) is utilized.

REFERENCES

1. The International Technology Roadmap for Semiconductors, SIA Handbook, 1998.
2. R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous and A. R. Leblanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” IEEE Journal of Solid-State Circuits, vol.9, pp.256-268, 1974.
3. A. Matsuzawa, "System Module with Chip on Chip Technology for Realizing True System LSI’s," in 8th Gakushin 165 committee meeting, pp.8-14, Nov. 2000.
4. K. Sasaki, S. Hanamura, K. Ueda, T. Oono, O. Minato, Y. Sasaki, S. Meguro, M. Tsunematsu, T. Masuhara, M. Kubotera, H. Toyoshima, “A 15-ns 1-Mbit CMOS SRAM,” IEEE J. Solid-State Circuits, vol. 23, pp. 1067- 1072, Oct. 1988.
5. M. Yoshimoto. K. Anami, H. Shinohara, T. Yoshihara, H. Takagi, S. Nagano, S. Kayano, and T. Nakano, “A Divided Word-Line Structure in the Static RAM and Its Application to a 64K Full CMOS RAM,” IEEE J. Solid-State Circuits, vol. 18, pp. 479- 485, May. 1983
6. A. Chandrakasan, W. Bowhill, and F. Fox, Design of High Performance Microprocessor Circuits. Piscataway, NJ: IEEE Press, 2000.
7. S. Mutoh, T. Douskei, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, “1-V power supply high-speed digital circuit technology with multi-threshold voltage CMOS,” IEEE J. Solid-State Circuits, vol. 30, pp. 847–854, Aug. 1995.