Ternary Content Addressable Memory

P.K.Sisira, N.Aswathy, B. Prameela, Anju George

Abstract*: *Memory Technology plays a vital role in fast searching applications. Content Addressable Memory (CAM) is a special type of memory used for search operation. CAM provides access to the stored data by its content instead of the address. Advanced version of CAM is known as Ternary CAM (TCAM) which is a memory that can also store don't care bit. TCAM is most relevant in routers in networking applications. Review of TCAM design techniques at different aspects are carried out, and obtained that an Energy Efficient TCAM (EE-TCAM) is the one which is having less power consumption. Compared with other SRAM-based TCAM designs, EE-TCAM use up reduced energy as it selectively activates only one row of SRAM at a time for search operation instead of activating the whole SRAM memory as in the other architectures. Partitioning of the TCAM table, designing of a pre-classifier and memory mapping are done prior to the work. This paper focuses on designing an EE-TCAM using Verilog HDL on Zybo7000 platform using Vivado design suite. Functional analysis of a 6*6 EE-TCAM is performed and power, delay and resource utilization are obtained. From the obtained results it is clear that EE-TCAM is having very less power and delay.

Keywords*: *Memory, TCAM, EE TCAM, Pre-classifier

I. INTRODUCTION

Content addressable memory (CAM) is also called as associative memory. CAM provides access to stored data by its contents instead of address and outputs the match address [1]. CAM compares the stored data with all the data stored in its memory to check if there is any match. If the data word is found, CAM returns the match addresses where the word was found. CAM has read and write operations similar to RAM. In addition to those operations, CAM supports fast data search, in a constant time. CAM is widely used in networking, neuromorphic associative memory, re-configurable computing, analytics, text mining, multimedia etc [2]. CAM is of two types, Binary Content Addressable Memory (BCAM) and Ternary Content Addressable Memory (TCAM). BCAM stores 1's and 0's, so that it requires exact match on the key-stored in its table. But TCAM allows partial match because it can store an additional don't care (x) bit along with 1's and 0's. The data searching flexibility is also high in case of TCAM compared to BCAM. Because of these advantages, TCAM is mostly used in most of the applications.

Revised Manuscript Received on May 20, 2020.

***** Correspondence Author

P. K. Sisira, Student, Department of Electronics & Communication, APJ Abdul Kalam Technological University. Kerala, India.

N. Aswathy, Assistant Professor, Department of Electronics & Communication, Adi Shankara Institute of Engineering & Technology, Kalady.

B. Prameela, Assistant Professor, Department of Electronics & Communication, Adi Shankara Institute of Engineering & Technology, Kalady.

Anju George, Assistant Professor, Department of Electronics & Communication, Adi Shankara Institute of Engineering & Technology, Kalady.

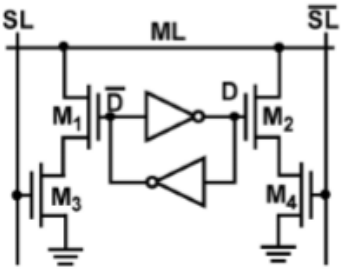

Fig 1. A BCAM Cell

A CAM cell performs two functions: bit storage and bit comparison. Bit storage portion consists of SRAM cell where cross coupled inverters are used for bit storage function. Bit comparison is done by using four NMOS transistors. TCAM cells are arranged as a two dimensional array. Cells in same column are connected by search lines(SL) and that of same row are connected by match lines(ML). Typical CAM cells (BCAM and TCAM) are illustrated in Figure 1 and in Figure 2. NMOS transistors M1, M2, M3, and M4 make comparison circuitry, D and Dbar are the bit storage nodes, ML represents the match line, and SL and SLbar are the search lines on which the input bit is to be applied in CAM cells. One SRAM cell is sufficient to store a bit in a BCAM cell and two cells are required to store a ternary bit in a TCAM cell. A typical BCAM cell has 10 transistors and a typical TCAM cell has 16 transistors [3].

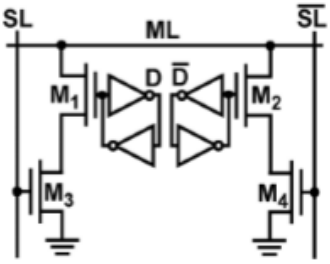

Fig 2. A TCAM Cell

A TCAM array is shown in Figure 3. For a search operation, all MLs are pre-charged to VDD and SLs are discharged to Ground. ML is constant at VDD only, if all the cells sharing the same ML have a match condition. A mismatch in any cell that shares same ML will discharge that ML to Gnd. The MLs are given to a priority encoder.

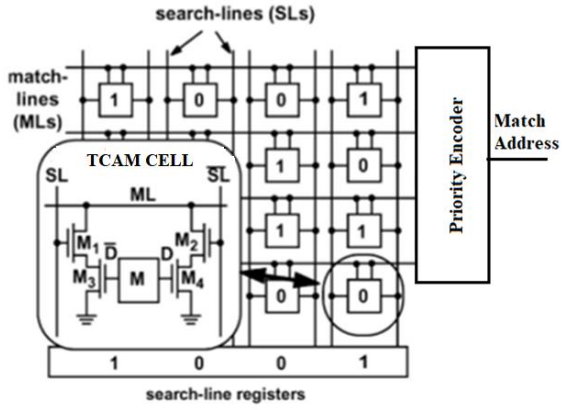

Fig 3. A TCAM Array

TCAM has certain limitations when compared to the Static RAM. It possesses size issues, issues of memory overflow and also it can't provide a deterministic operation. To overcome these issues Z. Ullah [3] proposed a new approach for TCAM design that is SRAM based TCAM. Classical TCAMs uses xor gate or NOR gate for comparison, but SRAM based TCAM uses SRAM for comparison.

II. BACKGROUND STUDY

The survey of SRAM based TCAM shows that there are only few papers in SRAM based TCAM. Parallel Hashing Memory proposed in [4] is a memory architecture which is an alternative to content addressable memory. It imitates CAM functionality with parallel RAM based hash tables. When the probability of failure of this memory architecture starts increasing before it reaches full capacity, the performance degradation occurs. Also it has a bucket overflow due to memory overhead. Sangyeun Cho [6] proposed a low power CAM on FPGA performs a hierarchical search on SRAM blocks. If a match is found in the SRAM block it stops the searching to reduce average power consumption. But worst case power consumption is high, which is a major concern in digital design.

Two types of FPGA implementations are presented in Xilinx application note [8]. First one is a CAM which is implemented using Xilinx BRAM resource and another one is a TCAM which is made up of 16-bit shift registers. The shift register-based TCAM works efficiently for smaller TCAMs.

SRAM-based TCAMs [3] developed to get higher throughput, appropriate partitioning methodology, better memory utilization, and support for large bit pattern and also provide a deterministic search performance of one word. There are many SRAM based TCAM implementations. Z. Ullah proposed four types of SRAM based TCAMs. All of them are arranged as layers and having same functionality

Hybrid Partitioned (HP) TCAM [10] is the first SRAM based TCAM. In HP-TCAM the classical TCAM table is partitioned in a hybrid way to obtain TCAM sub-tables, which are then mapped to SRAM memory. Search operation in HP-TCAM involves two SRAM followed by a AND operation. Z-TCAM proposed in [11] is more efficient than HP-TCAM. Efficient (E)-TCAM [12] has less number of layers compared to Z-TCAM and HP-TCAM. Also E-TCAM is more efficient compared with the first two SRAM based TCAM implementations. Another type of memory is UE-TCAM [12], which is an ultra energy efficient SRAM based TCAM memory, only having two layers. It is having lower resources utilization, lower power consumption, lower latency and higher efficiency compared with HP-TCAM, Z-TCAM and E-TCAM.

Resource Efficient SRAM based TCAM (REST) proposed in [14] consume comparable memory resources. However, REST is highly memory efficient with inbuilt Virtual Block methodology. Power consumption and number of logic blocks are less. Scalable TCAM proposed in [15] has a Scalable and modular architecture with multiple optimizations. Its throughput is high but it is having high power consumption. TCAM design using multi pumping enabled multi ported SRAM [16] have efficient memory utilization but excessive power consumption.

Almost all SRAM based TCAMs are having any one of the following three partitions. First partition is horizontal partitioning where row wise partition of TCAM table takes place. Second one is vertical Partitioning, where column wise partition of TCAM table takes place. And the last one is hybrid partition where column wise partition of TCAM table followed by row wise partition. All these partitions are complex and having larger area requirement.

Table.I Comparison of Different SRAM based TCAMs

Architecture	Size	FPGA	Speed MHz	Throughput Mbit/s	Power mW
HP-TCAM [10]	512*36	Virtex 6	35	4.2	188
Z-TCAM [11]	512*36	Virtex 6	118	5.6	109
E-TCAM [12]	512*36	Virtex 6	159	5.8	91
UE-TCAM [13]	512*36	Virtex 6	164	7.1	78
REST [14]	72*28	Virtex 6	35	1.4	161
Scalable TCAM [15]	1024*150	Virtex 6	20.4	3221	4587
EE-TCAM [16]	512*36	Virtex 6	336	11.6	33.7

From Table I , it is clear that Energy Efficient TCAM (EE-TCAM) is most efficient compared to other SRAM based TCAMs. EE-TCAM is introduced by Z.Ullah in 2018. This approach reduces the effect of partitioning in data search, using a pre-classifier based approach [16]. This architecture involves two classifications in design. First the TCAM table is divided into several sub-tables. The second SRAM based implementation stage maps the resultant TCAM sub-tables into SRAM blocks in the architecture. All other existing TCAMs energizes the entire SRAM memory for searching each incoming TCAM word. This system activates a part of SRAM memory for lookup instead of activating the entire SRAM memory, so power consumption is less compared to other architectures.

III. ARCHITECTURE AND CLASSIFICATION OF EE-TCAM

This paper focuses on designing a 6*6 EE-TCAM which follows a pre-classifier based approach. This architecture selectively activates at most one row of SRAM blocks for each incoming TCAM word. The TCAM is having some partitions in its table based on bits

A. *ARCHITECTURE*

The architecture is shown in Figure 4 which involves a pre-classifier unit and SRAM cells. The incoming W-bit TCAM word is sub divided into V sub-words. The sub-words are given as addresses to the selected row of SRAM blocks and corresponding SRAM words are read. The outputs of the SRAM units in each row are bit-wise ANDed together to obtain the resultant matching words and are provided to the associated priority encoder (PE). A Multiplexer is used at the output to obtain which priority encoder is activated.

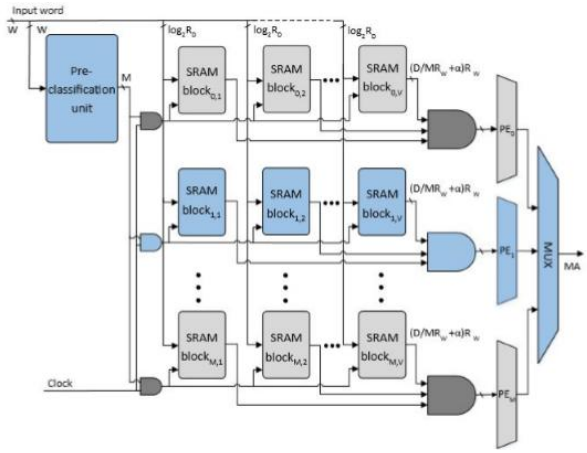

Fig 4. Architecture of EE-TCAM [16]

A 6*6 TCAM requires 4 rows of 2*4 SRAMs in its architecture. Each row consists of 2 SRAMs. These numbers of SRAM's are obtained by partitioning of the 6*6 TCAM table shown in Table II .

Table.II A 6*6 TCAM TABLE

Addresses	TCAM Words
0	001001
1	11x100
2	x10010
3	100x11
4	0x01x1
5	x10001

The TCAM words in Table 2 is partitioned into 4 sub tables based on b1 and b3 bits and stored in SRAM blocks by memory mapping. The partitioning and memory mapping details are provided in the classification section.

Figure 5 illustrates the Pre-classifier unit. This unit consists of multiplexers and address decoder in its architecture. The number of multiplexers depends on the partitioned size of EE-TCAM.

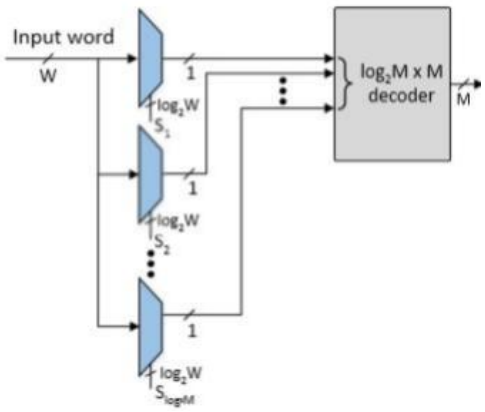

Fig 5. A Pre-Classifier unit [16]

Because there are 6 inputs, 3 select lines are needed for each of them. Because the partitions are based on b1 and b3 the select lines are 001 and 011. The output of multiplexer is given to a 2*4 decoder.

B. *CLASSIFICATION*

The partitioned 6*6 TCAM table is shown in Table III . The partitioning is based on b1 and b3 bits. If b1 or b3 are x, it is first represented as 0 and 1 before partitioning. First sub group (ST0) consists of TCAM addresses with b1=0 and b3=0. Next subgroups (ST1, ST2 and ST3) contain addresses with b1 and b3 as 01, 10 and 11 respectively. The partitioned addresses are mapped into the SRAM memory based on b0, b2, b4 and b5 bits. By this way, the partitioned TCAM table is mapped into 4 rows of SRAM block and each row consists of two 4*2 RAMs. The memory mapping represents how each word of TCAM is mapped to the memory. Each bit in TCAM table is mapped to the SRAM based on the table 2 below. Mapping includes populating the RAM, so that RAM based implementation can fulfill the same search function as TCAM. A TCAM cell is having two SRAMs in its architecture. So that a bit inside the TCAM table is mapped into two SRAMs for mapping. Each column of the RAM represents the match vector for a word.

Table.III Partition based on b1 and b3 bits

Addresses	TCAM words						
	b_0	b_1	b_2	b_3	b_4	b_5	
0	0	0	1	0	0	1	ST_0
3	1	0	0	0	1	1	
3	1	0	0	1	1	1	ST_1
4	0	0	0	1	x	1	
2	x	1	0	0	1	0	ST_2
5	x	1	0	0	0	1	
1	1	1	x	1	0	0	ST_3
4	0	1	0	1	x	1	

A wider TCAM deals with a wider input key. While implementing the TCAM in a single RAM, a wider input key is required, this is used as the address to access the RAM.

Table.IV Representing bits of TCAM in RAM

The value of the ternary bit	The value stored at RAM[0]	RAM[1]
0	1	0
1	0	1
don't care	1	1

Table IV represents bits of TCAM in RAM. The 0 in TCAM table is mapped by storing a 1 in RAM [0] and 0 in RAM [1]. The 1 in TCAM table is mapped by storing a 0 in RAM [0] and 1 in RAM [1]. The x (don't care) in TCAM table is mapped to by storing a 1 in both RAM [0] and RAM [1].

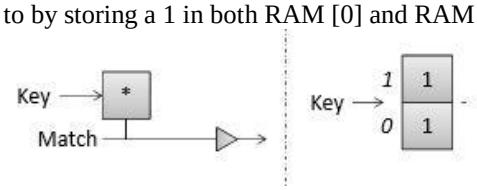

Fig.6 Mapping one bit in RAM [15]

Figure 6 shows matching a 1bit key with 1*1 TCAM. Here the key is *(don't care bit), which can be mapped to RAM in a way that both memory location 0 and memory location 1 can store '1' (From table 5).

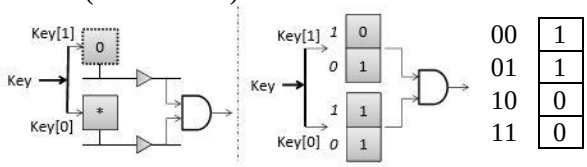

Fig.7 Memory Mapping of two bits [15]

Figure 7 shows mapping of two keys into a 1*2 RAM. Here the keys are '0' and don't care (*) bit. To map these keys corresponding value of each memory location are ANDed together. Figure 8 shows how the bits inside Table 3 are mapped to SRAM. Mapping takes place based on b1 and b3.

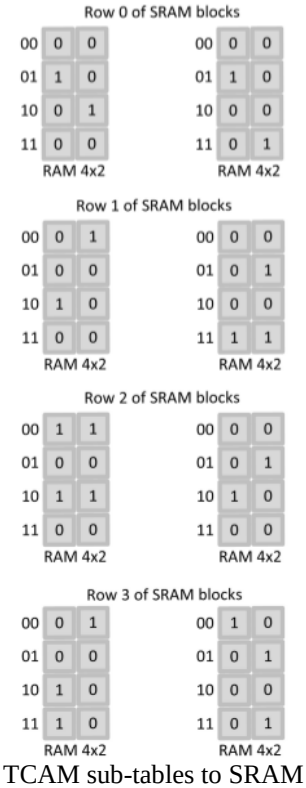

Fig.8 Mapping TCAM sub-tables to SRAM {b1, b3} [16]

First 2*4 RAM in each row is obtained by mapping b0 and b2 bits of each sub table in table 4 using the concept explained in Figure 7 and second RAM is obtained by mapping b4 and b5 of each sub table. The input data is searched against the bits inside these SRAMs,

IV. RESULTS AND PERFORMANCE EVALUATION

The output of a 6*6 EE-TCAM is shown in Figure 9. By pre-classification the memory location of stored word is obtained. Then the match address is obtained by bitwise AND operation. Here the TCAM table is represented as mem and mem1 to mem4 represent partitioned table. Mat1 to mat8 are dual port SRAM blocks. The data word (in) which is to be searched is given to the input of pre-classifier. The pre-classifier output (m-the output of address decoder inside the pre-classifier) is provided to SRAM blocks and search operation took place. At the input of priority encoder the memory location (memory) of the input word is obtained.

Each memory has two addresses; the priority encoder chooses the exact address of the given word. Since there is separated priority encoder for each row, the multiplexer chooses which priority encoder is activated. The final output (fin_out) is obtained at the multiplexer output.

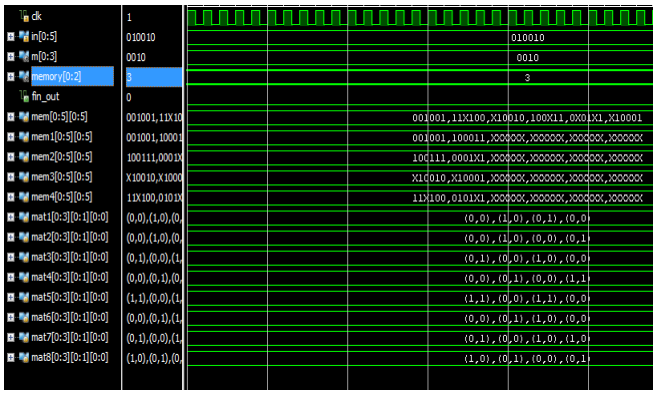

Fig.9 Final output of EE-TCAM

In this work, the pre-classifier is fed with an input 010010. The output of pre-classifier becomes 0100. Then after searching it is obtained that the input 010010 is on 0^{th} address of memory 3 .

Table.V Zynq-7000 FPGA resource utilization of 6*6 EE-TCAM

Resource	Utilization	Utilization %
LUT	440	0.20
LUTRAM	24	0.03
Flip Flop	237	0.05
IO	6	2.40

The proposed EE-TCAM designs were implemented in Zybo-7000 using Vivado HLS and power, area and delay are calculated. Table V lists the FPGA resource utilization parameters such as slice LUTs, LUTRAMs, Flip Flops and IO of the EE-TCAM design for 6*6 EE-TCAM.

Table VI represents the total power obtained after implementation and the delay

to obtain the output. The power is very less compared to other existing SRAM a based TCAM design in Table Ⅰ .

.Table.Ⅵ Power and Delay

Total Power (mW)	1.03
Dynamic Power (mW)	0.80
Static Power (mW)	0.231
Delay (ns)	4.237

Delay is also reduced compared to other designs. Speed of the system in MHz is obtained as 268MHz, which is high compared to other TCAMs in Table Ⅰ .

V. CONCLUSIONS

TCAMs are most relevant in networking routers. They can handle the demand of networking speed even as routing table and its size grows. There are Several SRAM based TCAMs, but most of them suffer from high power and resource utilization. Energy Efficient TCAM is the more power efficient TCAM compared to other SRAM based TCAMs. EE-TCAM uses pre-classification architecture which activates only one row of SRAM block for searching instead of activating entire SRAM memory as in other architectures. In this paper a 6*6 TCAM is designed using Verilog HDL on Zybo-7000 in Vivado HLS software. Partitioning of the TCAM table, designing of a pre-classifier and memory mapping are performed as a first step. After designing the complete system, the power, delay and resource utilization is obtained. For further optimizing the area a virtual box can be included in the system architecture.

.

REFERENCES

[1] K. Pagiamtzis and A. Sheikholeslami, Content-Addressable Memory (CAM) Circuits and Architectures: A Tutorial and Survey, IEEE Journal of Solid State Circuits, vol. 41, no. 3, pp.712-727, March 2006.
[2] Zahid Ullah, Manish Kumar Jaiswal, Y.C. Chan,and Ray C.C. Cheung, “FPGA Implementation of SRAM-based Ternary Content Addressable Memory”, 26th International Parallel and Distributed Processing Symposium Workshops PhD Forum,2012
[3] Ullah Zahid, “SRAM-Based Ternary Content Addressable Memory”, Doctor of Philosophy City University of Hong Kong, August 2014 .
[4] Mahoney, P., Savaria. Y., Bois, G., and Plante, P., “Parallel Hashing Memories: an Alternative to Content Addressable Memories”, 3rd International IEEE NEWCAS Conference, June 2005.
[5] Madian Somasundaram, “Circuit to Generate a Sequential Index for an Input Number in a Pre-defined List of Numbers”, US Patent, No. US 7155563 B1, December 2006.
[6] Sangyeun Cho, Martin, J. R, Ruibin Xu, Hammoud, M. H, and Melhem, R, “CA -RAM: A High-Performance Memory Substrate for Search-Intensive Applications”, proceedings of Performance Analysis of Systems Software, ISPASS, IEEE International Symposium, April, 2007.
[7] Ullah, Zahid. “LH-CAM: Logic-based higher performance binary CAM architecture on FPGA.” Embedded Systems Letters 9.2,2017
[8] Locke, K. Parameterizable content-addressable memory. In Xilinx Application Note XAPP1151; Xilinx: San Jose, CA, USA, 2011.
[9] Zahid Ullah and Sanghyeon Baeg, “Vertically Partitioned SRAM-based Ternary Content Addressable Memory,” Elsevier Journal, Procedia Engineering , 2011
[10] Ullah, Z. Ilgon, K. Baeg, ”Hybrid partitioned SRAM-based ternary content addressable memory.” IEEE Trans. Circuits Syst. I Regul., 2012
[11]Ullah, Z,Jaiswal, M.K. Cheung, R.C. “Z-TCAM: An SRAM-based architecture for TCAM.”, IEEE Trans. Very Large Scale Integr. Syst., 2015
[12] Ullah, Z.,Jaiswal, M.,Cheung, R., “E-TCAM: An Efficient SRAM-Based Architecture for TCAM” Syst. Signal Process, 2014
[13] Ullah, Z.; Jaiswal, M.K.; Cheung, R.C.C.; So, H.K.H. “UE-TCAM: An ultra efficient SRAM-based TCAM” In Proceedings of the IEEE Region 10 Conference (TENCON) Macau, November 2015
[14] Jiang, W. “Scalable ternary content addressable memory implementation using FPGAs”. In Proceedings of the Ninth ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), October 2013.
[15] Ahmed, A.,Park, K. Baeg, S. “Resource Efficient SRAM-Based Ternary Content Addressable Memory”, IEEE Transaction on Very Large Scale (VLSI) Integration Systems, April 2016
[16] Inayat Ullah , Zahid Ullah and Jeong-A Lee,”EE- TCAM: An Energy-Effcient SRAM-Based TCAM on FPGA”,MDPI Journal of Electronics,2018.

AUTHORS PROFILE

P.K.Sisira final year M.Tech VLSI and Embedded Systems student in APJ Abdul Kalam Technological university. Kerala. She received her B.tech degree in Electronics and Communication from University of Calicut Kerala.

N.Aswathy working as Asst. Professor in the Department of Electronics & Communication at Adi Shankara Institute of Engineering & Technology, Kalady. She received her B.Tech in Electronics and Communications from CUSAT and completed her ME with second rank in VLSI Design from Anna University. She is having 9 years of teaching experience. She has published 12 national and international journals. She is a professional member of IEEE.

B.Prameela, working as Asst. Professor in the Department of Electronics & Communication at Adi Shankara Institute of Engineering & Technology, Kalady She received her B.Tech degree in electronics & communication engineering from M.G University. Her M.Tech degree is in VLSI and embedded systems from. CUSAT Currently she is a research scholar in division of electrical engineering, School of Engineering, Cochin University of Science and technology. She is having 16 years of teaching experience.

Anju George, working as Asst. Professor in the Department of Electronics & Communication at Adi Shankara Institute of Engineering & Technology, Kalady. She received her B.Tech in Electronics and Communications from MG University and completed her MTech in VLSI and Embedded System from MG University. She is having 5 years of teaching experience. She has published 6 national and International Journals.