



Design and Implementation of Ternary Memory Using FPGA

Bandi Lakshmi Piyanka¹, Palli Srinivas²
M.Tech Student¹, Associate Professor²

Department of Electronics & Communication Engineering
BVC College of Engineering, Rajahmundry, India

Abstract:

Ternary content addressable memory (TCAM) is a memory with some special characteristics. TCAM performs high speed parallel search operations and the operation done in single clock cycle. But TCAM having some limitations as compared with SRAM, which are low storage density, circuit complexity and slow access time. So, further we can move to TCAM with hybrid partition, as Z-TCAM. This paper proposes TCAM functionality with SRAM. Here hybrid partition of stored data in memory blocks is more important. Hybrid partition is main reason of shrinking the size of the memory and latency time. The main goal is to implement SRAM using TCAM (Existing method) and SRAM using Z-TCAM (Proposed method). Here we will compare the area and delay reports for these existing and proposed methods. The tool here we used is xilinx 14.2v and the language used for verifying proposed implementation is Verilog /VHDL.

Keywords: Ternary Content Addressable Memory, Spin Transfer Torque RAM, Hybrid Partitioning, Memory Search, Low power, VLSI.

I. INTRODUCTION

CAM stands for Content Addressable Memory which is a special type of memory used by Cisco switches. In the case of ordinary RAM the IOS uses a memory address to get the data stored at this memory location, while with CAM the IOS does the inverse. It uses the data and the CAM returns the address where the data is stored. Also the CAM is considered to be faster than the RAM since the CAM searches the entire memory in one operation. CAM tables provide only two results: 0 (true) or 1 (false). [8] TCAM stands for Ternary Content Addressable Memory is the capability extension of CAM which can match a third state, which is any value. This makes TCAM a very important component of Cisco Layer 3 switches and modern routers, since they can store their routing table in the TCAMs, allowing for very fast lookups, which is considerably better than routing tables stored in ordinary RAM. TCAM is a specialized CAM designed for rapid table lookups [8]. TCAM cell has two static random access memory (SRAM) cells and a comparison circuitry and provides three state: 0, 1, and x where x is a don't care state. The x state is always regarded as matched irrespective of the input bit. TCAM provides single clock lookup with constant search time which makes it suitable for applications such as network routers, data compression, real-time pattern matching in virus detection, and image processing. Furthermore, the cost of TCAM is about 30 times more per bit of storage than SRAM. RAM is available in a wider variety of sizes and flavors, is more generic and widely available, and enables to avoid the heavy licensing and royalty costs charged by some CAM vendors. CAM devices have very limited pattern capacity and also CAM technology does not evolve as fast as the RAM technology. In paper [1] TCAM is designed using SRAM which is called as Z-TCAM, because even though the TCAM table provides lookup of entire table in single clock it has various disadvantages when compared to SRAM. TCAM cells,

comparator's circuitry in add complexity Page 2031 to the TCAM architecture. The access time of TCAM is 3.3 times longer than the SRAM access time due to the massive parallelism [5]. Complex integration of memory and logic also makes TCAM testing very time consuming [3]. The cost of TCAM is also about 30 times more per bit of storage than SRAM [6]. But, the parameters such as area, delay and power can be further reduced by using STT RAM instead of SRAM [1] which has been proved in this work. With the potential advantages of SRAM over CAM, and feasibility of FPGA technology, we propose a memory architecture called Z-TCAM that emulates TCAM functionality with SRAM.

II. CONTENT ADDRESSIBLE MEMORY

A CAM is a special type of storage memory TCAMs are one level higher than CAM because they can search unknown bits also i.e. ternary states. The main role of ternary content addressable memory (TCAM) is to search input data against the pre-loaded data and output the comparison result which is then used to invoke a related entry from a conventional memory. A TCAM cell has a mask cell, data cell, and masking and comparison circuitry. Mask and data cells are typically implemented with SRAM. TCAM is an outgrowth of RAM, which became popular in the literature for its high speed search operation. The major application of TCAM is in IPv6. Other applications are in network routers, cache memory, ATM switches, Translation look-aside Buffers (TLB) in micro processors. The parity bit based TCAM design consists of the original data segment and an extra one-bit segment, derived from the actual data bits. We only obtain the parity bit, i.e., odd or even number of "1"s. The obtained parity bit is placed directly to the corresponding word. The input to the structure is through search lines and the input is a search word. The size of the input word can be varied depends on the FPGA's size capacity. Here

we are using 8-bit input search words. Each stored word has a match line and this match line indicates the absence or presence of the search word inside the stored data. An encoder is used at the output of the CAM architecture to choose the output if multiple matches are detected. The encoder selects the output with the priority level.

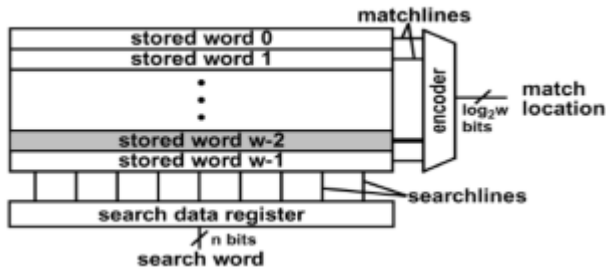


Figure.1. Basics structure of cam

While designing a new architecture our prime aim is to create it efficiently and provides maximum performance. The primary step for improvising the performance of TCAM is the hybrid partitioning. Hybrid partitioning logically dissects the TCAM table horizontally and vertically into $m \times n$ number of sub-tables. All TCAM sub-tables are then processed to be stored in their corresponding SRAM memory units. A conceptual view of hybrid partitioning is shown in the fig.2. Vertical partitioning part of hybrid partitioning implies that a TCAM word of width “W” bits are divided into “n” sub words, each of which is of width “w” bits. Horizontal partitioning part of hybrid partitioning divides each vertical partition using the original address range of conventional TCAM table. Hence, the dimension of each hybrid partition is “K w” where “K” represents a sub-set of original address pool and “w” is the number of bits in a sub-word. All TCAM tables have the same dimensions. Hybrid partitions/TCAM sub tables spanning the same address range is considered to be in the same layer. For example, $HP_{11}, HP_{12}, HP_{13}, \dots, HP_{1n}$ span the same address range and are in the same layer. It should be noted that number of layers are equal to number of horizontal partitions. As there are “m” horizontal partitions, thus there are “m” layers.

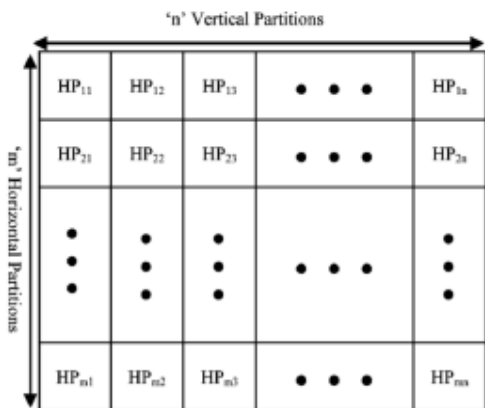


Figure.2. Hybrid partitioning

III. BASIC STRUCTURE OF CAM

The input to the structure is through search lines and the input is a search word. The size of the input word can be varied depends on the FPGA’s size capacity. Here we are using 8-bit input search words. Each stored word has a match line and this match

line indicates the absence or presence of the search word inside the stored data. An encoder is used at the output of the CAM architecture to choose the output if multiple matches are detected. The encoder selects the output with the priority level.

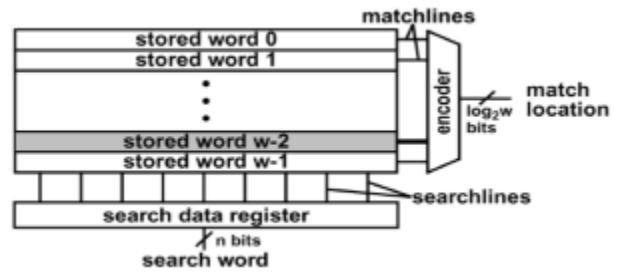


Figure.3. Allocator Structure of CAM

IV. MEMORY ARCHITECTURE OF SRAM BASED TCAM

Architecture of HP SRAM-based TCAM is depicted in Fig. 3 where each layer corresponds to Fig. 4. The output of each layer is a Potential Matching Address (PMA). In case of multiple PMAs (multiple matches), Global Priority encoder (GPE) selects the highest priority PMA as a Matching Address (MA). PMA of a lower layer has the highest priority. For example, if we have PMAs 1, 5, 9 corresponding to layers 1, 5, and 9 respectively, then GPE will select 1 as MA because it has the highest priority. Architecture of a layer of the proposed TCAM is shown in Fig. 4. Main components in a layer of the target memory architecture include “n” Bit Position Tables (BPTs), “n” Address position Tables (APT), “n” Address Position Table Address Generators (APTAGs), Local Priority Encoder (LPE), and ANDing operation. BPTs and APTs are constructed from SRAM. Each hybrid partition has its corresponding BPT, APTAG, APT, and ANDing operation.

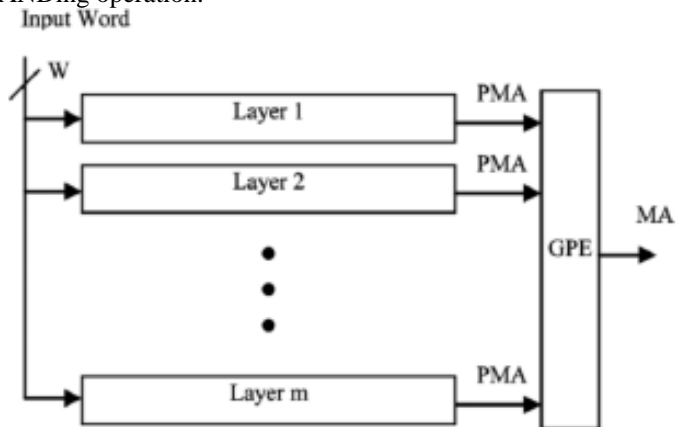


Figure.4. Memory Architecture of SRAM based TCAM

ARCHITECTURE OF Z-TCAM

• Overall Architecture

The overall architecture of Z-TCAM is depicted in Fig. 1 where each layer represents the architecture shown in Fig. 2. It has L layers and a CAM priority encoder (CPE). Each layer outputs a potential match address (PMA). The PMAs are fed to CPE, which selects match address (MA) among PMAs.

• Layer Architecture Layer architecture is shown in Fig. 2. It contains N validation memories (VMs), 1-bit AND operation, N original address table address memories (OATAMs), N original

address tables (OATs), K-bit AND operation, and a layer priority encoder (LPE).

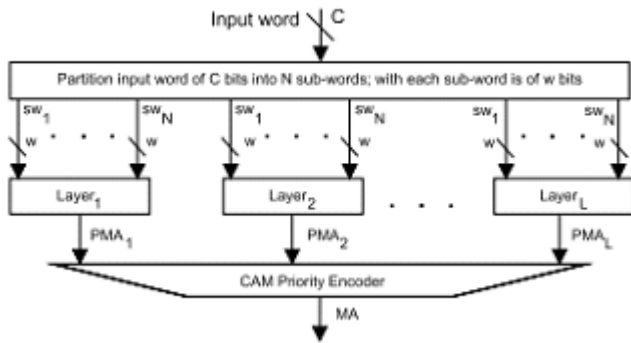


Figure.5. Architecture of Z-TCAM

V. IMPLEMENTATION RESULTS

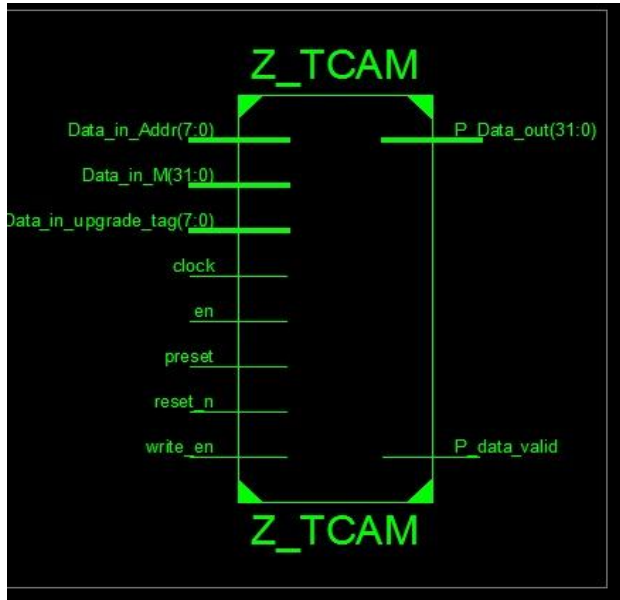


Figure.6. block diagram of Z TCAM

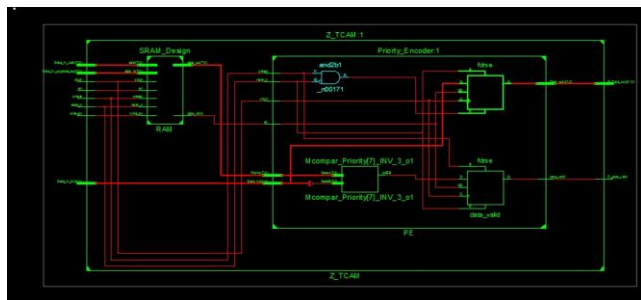


Figure.7. RTL Schematic for Z TCM

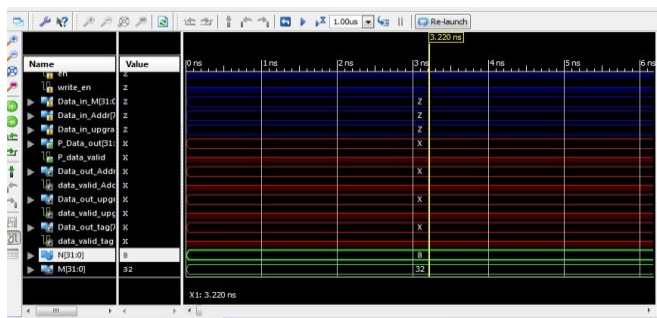


Figure. 8. Simulation for Z TCAM

The following are the existing and proposed area and delay synthesis reports as shown below:

AREA:

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	42	18224	0%
Number of Slice LUTs	79	9112	0%
Number of fully used LUT-FF pairs	42	79	53%
Number of bonded IOBs	86	232	37%
Number of BUFG/BUFGCTRLs	1	16	6%

Figure.9. Existing Method

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	60	18224	0%
Number of Slice LUTs	97	9112	1%
Number of fully used LUT-FF pairs	60	97	61%
Number of bonded IOBs	86	232	37%
Number of BUFG/BUFGCTRLs	1	16	6%

Figure.10. Proposed Method

DELAY:

Offset: 4.239ns (Levels of Logic = 2)
 Source: reset_n (PAD)
 Destination: RAM/Mram_SRAM7 (RAM)
 Destination Clock: clock rising

Data Path: reset_n to RAM/Mram_SRAM7

Cell:in->out	fanout	Gate Delay	Net Delay	Logical Name (Net Name)
IBUF:I->O	83	1.222	2.013	reset_n_IBUF (reset_n_IBUF)
LUT4:I0->O	8	0.203	0.802	RAM/Mmux_BUS_000111 (RAM/BUS_0001)
RAM256X1S:WE		0.000		RAM/Mram_SRAM2
Total				4.239ns (1.425ns logic, 2.814ns route) (33.6% logic, 66.4% route)

Figure.11. Existing Method

Delay: 2.612ns (Levels of Logic = 2)
 Source: RAM/data_out_5 (FF)
 Destination: PE/data_valid (FF)
 Source Clock: clock rising
 Destination Clock: clock rising

Data Path: RAM/data_out_5 to PE/data_valid

Cell:in->out	fanout	Gate Delay	Net Delay	Logical Name (Net Name)
FDS:C->Q	2	0.447	0.845	RAM/data_out_5 (RAM/data_out_5)
LUT6:I3->O	1	0.205	0.808	PE/data_valid_glue_set_SWO (N7)
LUT6:I3->O	1	0.205	0.000	PE/data_valid_glue_set (PE/data_valid)
FDR:D		0.102		PE/data_valid
Total				2.612ns (0.959ns logic, 1.653ns route) (36.7% logic, 63.3% route)

Figure.12. Proposed Method

VI.CONCLUSION

In this brief, we have presented a novel SRAM-based TCAM architecture of Z-TCAM. We have implemented two example designs of 512 × 36 and 64 × 32 of Z-TCAM on Xilinx spartan 6 FPGA. FPGA implementation is a big plus for Z-TCAM. Resources utilization, speed, and power consumption for different situations for the example designs on FPGA as well as in ASIC have been tabulated. Z-TCAM also ensures large capacity TCAM whereas this capability is lacked by

conventional ones. Moreover, the proposed TCAM has a simpler structure, and very importantly, has a deterministic search performance of one word comparison per clock cycle

mechanism for fast table lookup,” U.S. Patent 20 060 253 648, Nov. 2, 2006.

VII. REFERENCES

[1]. N. Mohan, W. Fung, D. Wright, and M. Sachdev, “Design techniques and test methodology for low-power TCAMs,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 6, pp. 573–586, Jun. 2006.

[2]. P. Mahoney, Y. Savaria, G. Bois, and P. Plante, “Parallel hashing memories: An alternative to content addressable memories,” in *Proc. 3rd Int. IEEE-NEWCAS Conf.*, Jun. 2005, pp. 223–226.

[3]. S. Dharmapurikar, P. Krishnamurthy, and D. Taylor, “Longest prefix matching using bloom filters,” *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 397–409, Apr. 2006.

[4]. D. E. Taylor, “Survey and taxonomy of packet classification techniques,” *ACM Comput. Surveys*, New York, NY, USA: Tech. Rep. WUCSE-2004-24, 2004.

[5]. P. Mahoney, Y. Savaria, G. Bois, and P. Plante, “Transactions on high-performance embedded architectures and compilers II,” in *Performance Characterization for the Implementation of Content Addressable Memories Based on Parallel Hashing Memories*, P. Stenström, Ed. Berlin, Germany: Springer-Verlag, 2009, pp. 307–325.

[6]. S. V. Kartalopoulos, “RAM-based associative content-addressable memory device, method of operation thereof and ATM communication switching system employing the same,” U.S. Patent 6 097 724, Aug. 1, 2000.

[7]. W. Jiang and V. Prasanna, “Scalable packet classification on FPGA,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 9, pp. 1668–1680, Sep. 2012

[8]. M. Becchi and P. Crowley, “Efficient regular expression evaluation: Theory to practice,” in *Proc. 4th ACM/IEEE Symp. Archit. Netw. Commun. Syst.*, Nov. 2008, pp. 50–59.

[9]. Xilinx, San Jose, CA, USA. Xilinx FPGAs [Online]. Available: <http://www.xilinx.com>

[10]. W. Jiang and V. K. Prasanna, “Large-scale wire-speed packet classification on FPGAs,” in *Proc. ACM/SIGDA Int. Symp. Field Program. Gate Arrays*, 2009, pp. 219–228.

[11]. W. Jiang and V. Prasanna, “Parallel IP lookup using multiple SRAMbased pipelines,” in *Proc. IEEE Int. Symp. Parallel Distrib. Process.*, Apr. 2008, pp. 1–14.

[12]. S. Cho, J. Martin, R. Xu, M. Hammoud, and R. Melhem, “CA-RAM: A high-performance memory substrate for search-intensive applications,” in *Proc. IEEE Int. Symp. Perform. Anal. Syst. Softw.*, Apr. 2007, pp. 230–241.

[13]. M. Somasundaram, “Memory and power efficient