

A 5nm Fin-FET 2G-search/s 512-entry x 220-bit TCAM with Single Cycle Entry Update Capability for Data Center ASICs

Chetan Deshpande, Ritesh Garg, Gajanan Jedhe, Gaurang Narvekar, Sushil Kumar
MediaTek Inc., San Jose CA, USA

Email: {Chetan.Deshpande, Ritesh.Garg, Gajanan.Jedhe, Gaurang.Narvekar, Sushil.Kumar}@mediatek.com

Abstract

This paper presents a 2G-search/s embedded Ternary Content Addressable Memory (TCAM) design in 5nm Fin-FET technology with the ability to update both SRAM words in a TCAM entry in a single clock cycle. This reduces TCAM update latency by 50% for data center Application Specific Integrated Circuits (ASICs) with only 1% area overhead and no search power penalty. We present a novel time multiplexed input bus interface on a single port TCAM cell array and new architecture to enable fast updates. Silicon measurement shows the highest reported search rate of 2G-search/s at a 3.48Mb/mm² memory density including all global peripheral circuitry for a 512 entry, 220-bit wide, 110Kb TCAM.

Introduction

TCAMs are used widely in network switches for packet classification and forwarding to perform fast lookups. Recently, due to the wide proliferation of Software Defined Networks (SDN), data center ASICs have to be more SDN compatible which requires support for quick rule updates to respond to dynamically changing network conditions. Each rule update in TCAM can trigger multiple entries to be moved since rules are stored in order of priority in the memory array. Under these conditions, TCAM update latency becomes a major bottleneck [1]. Each TCAM entry update requires two spatially adjacent SRAM memory words to be written, typically referred to as DATA (X) and MASK (Y) words (Fig.1). Together, data stored in these two words facilitates the ternary function in TCAM (Fig.1 (b)). Therefore, at line rate with a typical single port TCAM interface, TCAM entry update takes two clock cycles. Further, frequently the system cannot accommodate two back-to-back clock cycles to complete a TCAM update, sometimes having to wait several thousand clock cycles after partially updating an entry. In the intervening cycles, packets can get misdirected or dropped. In this paper, we propose a novel interface and architecture to facilitate one cycle TCAM update at 2GHz clock frequency with 2G-search/s performance.

Proposed Design and Architecture

A typical TCAM input bus interface is shown in Fig.2. SDI[n-1:0] and MASKB[n-1:0] supply the search and column masking key (Fig.3) during a search operation respectively. SDI[n-1:0] also drives the memory write data during a write operation. The timing diagram in Fig.4 (a) illustrates the typical write operation taking two clock cycles to update the TCAM entry. A[m:0] bus supplies the address of the memory location to be updated. SDI[n-1:0] drives data word (X) on cycle 1 and mask word (Y) on cycle 2 to fully update the TCAM entry. In order to do a Single Cycle Update (SCU) to TCAM entry, X and Y word data have to be supplied on the same cycle. One way to solve this is by adding another input port and bus to supply the Y-word. However, this increases the area and power overhead while also changing the user interface. In this work, we time multiplex the MASKB[n-1:0] bus and search circuitry to supply the Y-word during the SCU operation. A timing diagram with this user interface is shown

in Fig.4 (b). Here both X and Y words are supplied on SDI and MASKB respectively, in parallel. On the rising edge of clock both SDI and MASKB are captured and the lowest address bit A[0] is ignored for SCU operation. As shown in Fig.5, the MASKB bus is internally coupled to the SDI bus for column masking using logic gates 'C' and 'D' during a search operation to generate the complimentary signals SL and SLB going to the array for search. We introduce a new global control pin 'SCU' along with logic components 'A' and 'B' for each IO, as shown in Fig.5, to decouple the MASKB data from the SDI latch output during the SCU operation. In our design, when SCU is asserted, logic component 'B' ensures that column masking in 'C' and 'D' is disabled and logic component 'A' feeds the MASKB data carrying the Y-word to the SDI latch. Thus we reuse the search circuitry for SCU with only 1% area overhead. This scheme is fully backward compatible with conventional 2 cycle update. The output of the SDI and wrdata latch is served to a new 2-to-1 multiplexer which selects either the data on the internal write bus or the search bus depending on when the internally self-timed signal 'SCU_MUXSEL' is toggled.

The array architecture was modified to support 2 GHz search speed and also facilitate two memory writes in single cycle with sufficient write margin. We use a shallower bank of 128 TCAM entries in this design with the write driver, read sense amplifier and search line driver in the middle (Fig.6). We leverage the fact that in TCAM update, both memory write operations occur in spatially adjacent rows as shown in Fig.1, thereby eliminating the need for the second memory write address to setup before write can be performed. As shown in Fig.7, the array decoder is designed such that all word line drivers (WLDRV) on even rows and odd rows share enable signals XPZE and XPZO respectively. The xdecoder decodes the address to the lowest odd-even pair using pre-decoded lines from A[m:1]. Referring to Fig.4 (b), when SCU is asserted, at rising edge of clock, a self-timed signal XPZE triggers the X-word line. When XPZE resets after a programmable delay, it toggles SCU_MUXSEL to switch the data bus and also triggers a non-overlapping self-timed signal XPZO, triggering the Y-word line. XPZO and SCU_MUXSEL reset after a programmable delay completing the Y-word write and triggering the BL pre-charge. To avoid cycle time penalty there is no bit line pre-charge between the two writes since pre-charge is not needed between two back-to-back write operations. The critical path for the SCU operation is the data setup time with respect to the Y-word line. Our simulation shows a write margin >30% for the 6-sigma weak bitcell in the worst SFG/0.72V/-40C corner (Fig.8). No search power overhead is observed with SCU capability.

Measurement Results

This TCAM was designed in a 5nm Fin-FET technology. The test chip photo and layout plot are shown in Fig.9 (a) & (b). Fig.10 shows the Shmoo plot of our design on silicon. We demonstrate a 2 GHz SCU performance at 0.8V supply and room temperature. The highest reported search performance of

2G-search/s is achieved along with SCU operation on a 512-entry x 220-bit macro. Table 1 summarizes the comparison of this work with previously published papers.

Table 1 Comparison of this work with previously published work

Reference	[2]	[3]	[4]	[5]	This Work
Process	16nm	14nm	12nm	7nm	5nm
Entries (ML#)	128	256	128	256	512
Key Width (SL#)	80	160	80	80	220
Memory Density (Mb/mm ²)	1.80	2.01	1.08	4.04	3.48
Search Speed (search/s)	1.25G	1.40G	1.50G	1.60G	2.00G

References

- [1] K.Qiu, et al., IEEE ICDCS, pp.918-927, July 2018.
- [2] Y.Tsukamoto, et al., VLSI Cir. Dig., pp. C274-C275, June 2015.
- [3] I.Arsovski, et al., JSSC, vol. 53, no.1, pp. 155-163, Oct. 2017.
- [4] M.Yabuuchi, et al., VLSI Cir. Dig., pp. C19-C20, June 2018.
- [5] M.Yabuuchi, et al., VLSI Cir. Dig., CM1.5_192, June 2020.

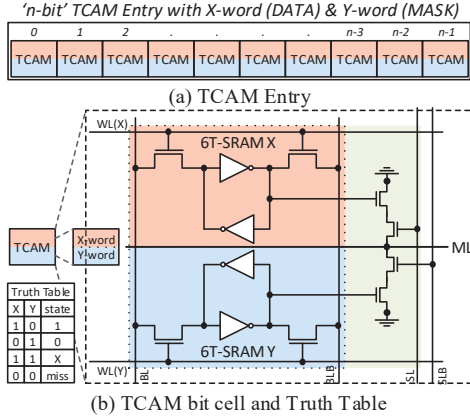


Fig. 1 TCAM Entry, bit cell and Truth Table

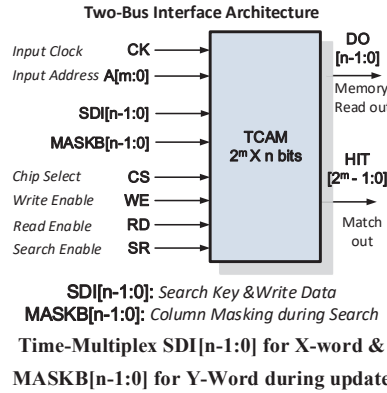


Fig. 2 Input block diagram

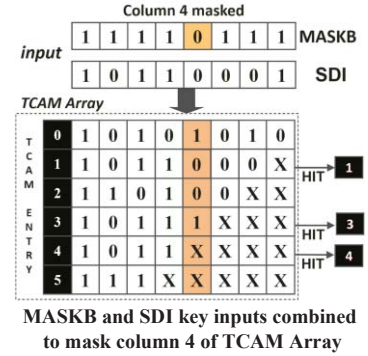


Fig. 3 Column masking

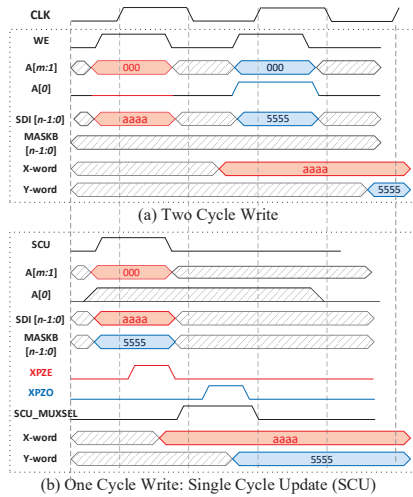


Fig. 4 Timing diagram

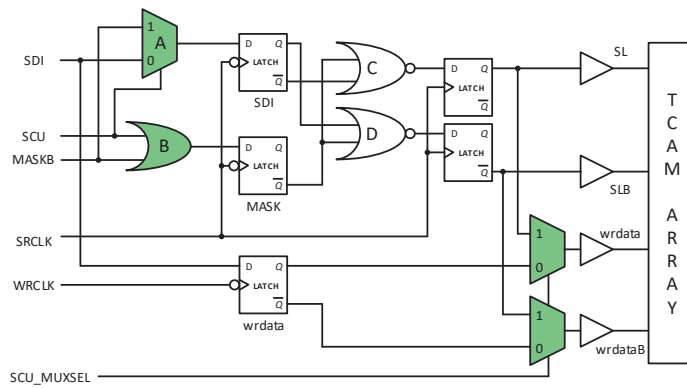


Fig. 5 Proposed Time-multiplexed IO interface for TCAM

MATCH LATCH	Control	MATCH LATCH
XDECODER		XDECODER
VALID BIT		VALID BIT
64 Entries X 110 bits	SLASLE Driver Write Driver Read Sense Amplifier	64 Entries X 110 bits
TCAM ARRAY		TCAM ARRAY

← 128 TCAM Entries →

Fig. 6 TCAM bank

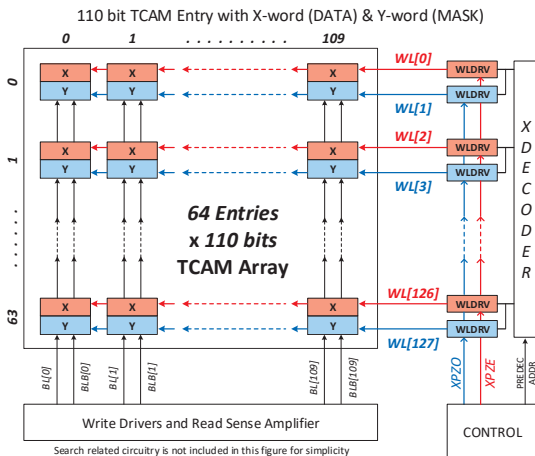


Fig. 7 TCAM Array with proposed decoding with odd/even row enable

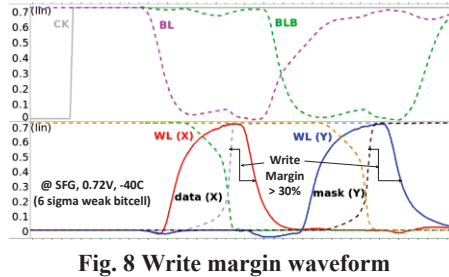


Fig. 8 Write margin waveform

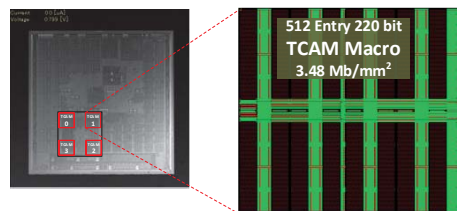


Fig. 9 (a) Die photo Fig. 9 (b) Layout plot

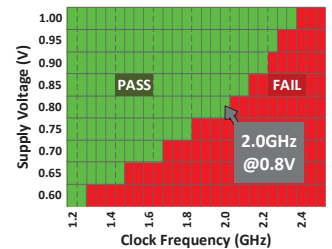


Fig. 10 Shmoo plot at Room temperature