SCP-TCAM: A Power-Efficient Search Engine for Fast IP Lookup

Xudong Li, Zhen Liu, Wei Li, Bin Liu
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, P.R.China
{lixd00, liuzhen02, li-wei03}@mails.tsinghua.edu.cn   liub@tsinghua.edu.cn

***Abstract*- We propose a power-efficient routing table search engine architecture, Stage-Control Pipelined TCAM (SCP-TCAM), for high-throughput router systems. SCP-TCAM can perform one IP lookup a clock cycle in a pipelined fashion while keeping the advantages of simple management and incremental updates of conventional TCAM. The most important feature of SCP-TCAM is that it consumes only 1/5 to 1/20 of the power compared to conventional TCAM of the same size. We also implement our proposed scheme using a Xilinx Field Programmable Gate Array (FPGA) to further evaluate the performance. SCP-TCAM will be a good alternative scheme for TCAM in high-performance, low-power consumption IP lookup engine.**

I. Introduction

Due to the explosive growth of the Internet, Classless Inter-Domain Routing (CIDR) was widely adopted to prolong the life of Internet Protocol Version 4 (IPv4) [9]. Ever since the introduction of CIDR, the IP address lookup mechanism has been designed on the base of the longest prefix matching (LPM) algorithm, which requires Internet routers to search variable-length address prefixes in order to find the longest matching prefix of the IP destination address and retrieve the corresponding forwarding information for each packet traversing the router. IP address Lookup is often the performance bottleneck in high-performance Internet routers.

Many techniques are available to perform IP address lookups. One of the most commonly used approaches is employing Ternary Content Addressable Memory (TCAM) devices. However, TCAMs have traditionally been more expensive and consumed more power than ASIC-based solutions. While the cost to density ratio keeps decreasing, the high power consumption becomes a significant factor that blocks TCAM from being deployed in high-end routing systems. An 18Mbit TCAM typically consumes as much as 12-15 Watts when all entries are enabled for search.

Several mechanisms have been proposed to reduce power consumption by selectively addressing smaller portions of the TCAM. Each portion (called a sub-table or database) is defined as a set of TCAM blocks. A TCAM block is a contiguous, fixed-sized chunk of TCAM entries, usually much smaller than the size of the entire TCAM [16]. However, these methods complicate routing table management and the power reduction is constrained by partitioning algorithm and the characteristics of specific routing table.

In this paper, we present a novel hardware architecture called SCP-TCAM (stage-controlled pipeline TCAM) for fast IP address lookup. SCP-TCAM can achieve a power reduction factor of 5 to 20 compared to TCAM of the same size, which makes it a promising scheme for high-end routing systems.

The rest of the paper is organized as follows. Section II summarizes related work. The architecture of SCP-TCAM is proposed in section III. The simulation result and a FPGA implementation is presented in section IV. A conclusion is drawn in section V.

II. Previous Work

IP address lookup has been an active area of research in recent years. Some schemes for fast IP lookup algorithms [1]–[19] have been proposed in the literature.

The available IP address lookup schemes can be generally classified into two categories, binary search-based algorithms and TCAM-based routing tables. The former include various trie-based approaches [9][11], prefix-based binary search [5][13], multiway and multicolumn search [8], prefix expansion [6], and so on. They can be implemented in software or hardware, but they typically incorporate sophisticated data structure and are hard to manage and update. On the contrary, TCAM-based solutions are simple to manage and able to provide deterministic search latency. It has been widely investigated because of its being a promising solution for high performance routing engines.

Recently, several novel ideas have been introduced to further improve the performance of various search schemes. Sharma et al. in [14] proposed some sorting and searching algorithms using TCAM, which can be finished in O(n) memory cycles. Two techniques compacting routing tables were presented by Liu Huan in [12]. Desai and Gupta [15] implemented a novel architecture for IP address lookup using a large programmable Finite State Machine (FSM). The algorithm of Bloom filters is employed for Longest Prefix Matching (LPM) in [17]. Hyesook et al. [18] solved the LPM problem by converting it into an exact matching problem. Work is also done to reduce the power consumption of TCAM. Zane and Narlikar [16] have developed several algorithms called CoolCAMs to make the routing engine power efficient.

III. ARCHITECTURE OF SCP-TCAM

A TCAM (Ternary Content Addressable Memory) stores data with three logic values: ‘0’, ‘1’ or ‘X’ (don’t care). Each entry contains a value bit-string and mask bit-string. For a given input, the TCAM compares it against all of the entries in parallel, and returns the first entry that matches the input. When multiple entries match the search key, the entry with the highest priority

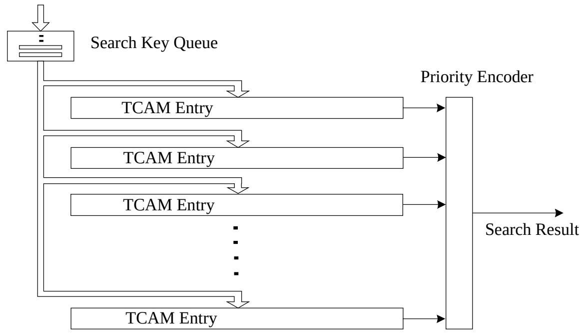

Fig. 1 Block diagram of conventional TCAM

is returned as the result and in most implementations, lower address usually has higher priority. Thus, a single TCAM access is sufficient to perform a route lookup operation. Moreover, TCAM-based tables are typically much easier to manage and update than tables using tries. The schematic diagram of TCAM is shown in Fig.1.

However, high hardware parallelism brings high power consumption, which keeps it from being employed in high-volume routing tables and high-end routing systems. SCP-TCAM is proposed aiming to solve this problem. In the following part, we will illustrate SCP-TCAM in detail.

Given that the power consumption of a TCAM is linearly proportional to the number of searched entries, we can further conclude that it is also linear proportional to the number of bits participating in the comparison if we can provide an efficient control mechanism. For example, the power consumption of a 4-bit lookup is only one eighth of that of a 32-bit one. In addition, an important but often overlooked fact is that the whole entry will not match the input as long as there is one unmatched bit. In other words, if we compare the corresponding bit one by one, once an unmatched bit is found, the other comparisons within that entry are no longer needed.

Based on the two observations above, we now present a new scheme, “Stage-Controlled Pipeline TCAM (SCP-TCAM)”, which results in a power reduction factor of 5~20 comparing with the existing TCAM solution in our experiments. Typical techniques such as “Parallel Comparison” and “Priority Encoder” used in conventional TCAMs are also employed in our solution. The key optimization is the “Stage-Controlled Pipeline” scheme, which substantially improves the efficiency of each entry lookup.

In a SCP-TCAM, the lookup operation in each routing table entry can be described as below: The first bit (leftmost bit) is compared with the first bit of the input at the first clock cycle. If a mismatch happens, a signal of “stop” will be sent to the second bit, and the whole comparison is terminated; if they are identical, a signal of “go on” will be sent to the next bit, which allows the comparison between the second bits of the search key and the entry. This procedure will continue until the last bit is compared. When all the bits are matched, the “Result” signal will be set high to indicate that. The corresponding bit-comparisons within different entries are performed in parallel just as in the conventional TCAM. Under this “Stage-Controlled” mechanism, whether one bit in the entry will be compared with the corresponding bit in input is controlled by the comparing result of its previous bits. In addition, because only one bit is compared in a clock cycle and pipeline mechanism is employed, we call the scheme, “Stage-Controlled Pipeline”.

Fig.2 illustrates the idea of SCP-TCAM. T1~Tn are the tri-state individual bits of a table entry stored in TCAM cells. B1~Bn are the corresponding bits in input IP address. As described above, difference between any two compared bits will terminate the comparison chain. If all bits are identical, the signal “Result” sent out by the last comparator will be set high.

We introduce “Power Diagram” to illustrate the distribution of searched bits that cause power consumption during lookup operation. The shaded bits

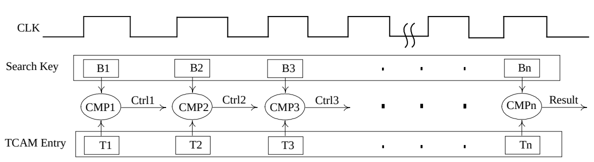

Fig.2 Comparison chain within an entry

in the “power diagram” are the ones that have to consume power. It is easy to obtain the total power consumption from the area of shaded bits in “power diagram”.

Fig 3 shows the “Power Diagram” of routing table entry 206.11.58.99/28 in a conventional TCAM when the input IP address is 10.43.75.196. Assume that the bits with the corresponding mask bits being 0’s do not take part in the comparisons, then the leftmost 28 bits of the entry will be compared with the input IP address for IP lookup since the prefix length is 28.

Fig 4 shows the “Power Diagram” of routing table entry 206.11.58.99/28 in a STP-TCAM when the input IP address is 10.43.75.196. Note that the entire search operation for this entry is terminated when the comparisons forward to the fourth bit because only the leftmost 3 bits are identical with the given input IP address.

As for the whole routing table, since there is a very small percentage of entries matching the input IP address and the majority of entries only match a few bits, the power consumption of the whole SCP-TCAM can be reduced significantly. The “Power Diagram” for the entire SCP-TCAM is almost blank. A possible distribution is shown in Fig. 5.

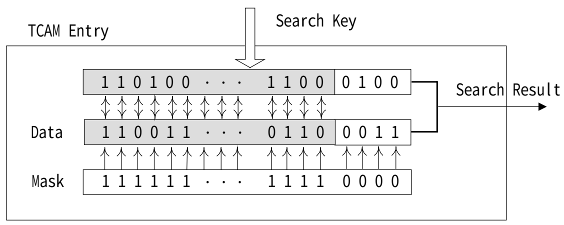

Fig.3 The “power diagram” of a typical TCAM

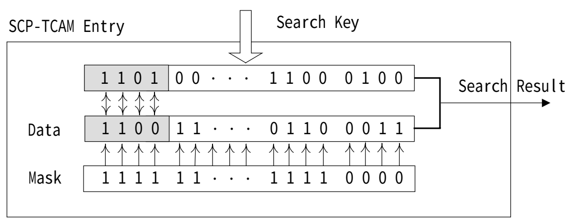

Fig.4 The “power diagram” of a SCP-TCAM

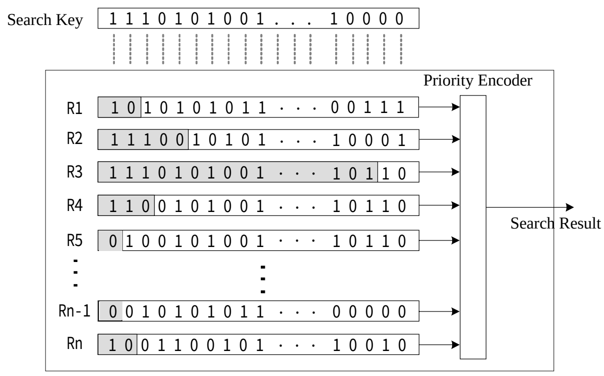

Fig.5 The "power diagram" of the whole routing table

IV. PERFORMANCE AND FPGA IMPLEMENTATION

As mentioned above, since the power consumption of a TCAM is linearly proportional to the number of searched bits of all routing table entries, we use this number to evaluate the power consumed by SCP-TCAM. Three synthetic routing tables from NPF (Network Processing Forum) are employed in our experiment, with sizes of 28K, 135K and 1M respectively. To ensure the generality of simulation results, we choose the IP addresses of four frequently visited websites, www.google.com (216.239.59.99), www.msn.com (207.68.173.244), www.stanford.edu (171.67.16.81) and www.tsinghua.edu.cn (166.111.4.100) as the given search keys.

Table 1 lists the ratios of the number of searched bits in SCP-TCAM and conventional TCAM for the four IP addresses in different routing tables.

From these data we can see that the ratio ranges from 5.3% to 20.2%. We can also find that it decreases as the routing table grows larger, which can be explained as that larger table introduces more irrelevant entries which can be aborted during comparison of the first few bits. As for the Telstra_synth_1M, all the ratios are well under 6.1%, in another word, a power reduction factor of more than 16 is achieved.

TABLE 1 THE PERCENTAGE OF POWER CONSUMPTION

Search Key	Routing Table		
	Mae-west	telstra_sanitized	telstra_synth_1M
Google	17.4%	15.9%	5.8%
Msn	20.2%	18.7%	6.1%
Stanford	9.5%	8.9%	5.4%
Tsinghua	9.7%	9.0%	5.3%

In order to further evaluate the proposed architecture, we develop a Field Programmable Gate Array (FPGA) implementation with a Xilinx xcv2000e. The internal block diagram is shown in Fig .6.

We connect the comparison result signal to the clock-enable port of register storing the next bit within an entry. Thus the power consumption is effectively controlled. The lower part of Fig. 6 shows these entries. In addition to the "Parallel Comparison", a "Pipeline Controller" is incorporated to ensure that the comparison progresses in a pipeline mechanism.

Let n denote the width of an entry within a SCP-TCAM. There are n n-bit Key Registers and n 1-bit Bit Registers in the Pipeline Controller, denoted as KREG and BREG for simplicity. They are used to hold the content of search key and the comparison bit of each stage, respectively. When a search key K[1..n] comes, it is latched into KREG[1] at the first clock cycle. At the second clock cycle, the content of KREG[1] is transferred to KREG[2] and the first bit of KREG[1] is latched into BREG[1] for comparison with the first bits of all entries in the SCP-TCAM. The content of K[1..n] continues to flow through all the KREGs during the

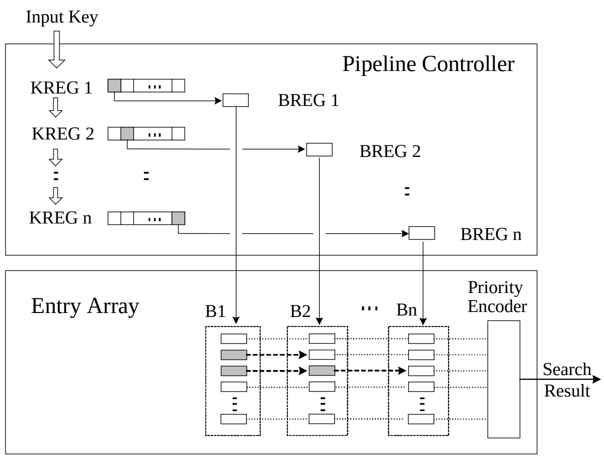

Fig.6 A hardware model of SCP-TCAM

following cycles and the corresponding bits are sent to BREGs consecutively until we get the final result. This process can work in a pipeline fashion and provide one searching result every cycle.

We implemented one hundred 32-bit entries, which cost 1581 slices and can work at 170MHz. Because we only compare one bit during one clock cycle, SCP-TCAM can easily achieve a higher frequency than conventional TCAM.

V. CONCLUSION

IP address lookup is one of the primary functions of the router and is a major performance bottleneck. TCAM is one of the most attractive solutions for high-throughput routing systems. However, in spite of the advantages of fast lookup and easy management, it also suffers from high power consumption. Several mechanisms have been proposed to reduce power consumption by selectively addressing smaller portions of the TCAM. But these methods complicate the construction and update of forwarding table and the power reduction is constrained by the number of blocks contained in the TCAM devices.

In this paper, we present a new search engine architecture called SCP-TCAM that effectively reduces the power consumption while still offering ultra-high search capability. Our simulation results show that a power reduction factor of 5 to 20 can be achieved. A FPGA implementation is also developed to further evaluate our scheme.

ACKNOWLEDGMENT

The authors would like to acknowledge support received from members of Broadband Network Switching Technology and Communication Lab, Department of Computer Science and Technology, Tsinghua University. The authors would also like to thank the anonymous referees for their valuable feedback on this work.

REFERENCES

[1] Pei, T.-B.; Zukowski, C., "VLSI implementation of routing tables: tries and CAMs", INFOCOM '91. Proceedings. Tenth Annual Joint Conference of the IEEE Computer and Communications Societies. Networking in the 90s. IEEE, 7-11 April 1991 Page(s): 515 -524 vol.2

[2] McAuley, A.J.; Francis, P., "Fast routing table lookup using CAMs", INFOCOM '93. Proceedings. Twelfth Annual Joint Conference of the IEEE Computer and Communications Societies. Networking: Foundation for the Future. IEEE, 28 March-1 April 1993 Page(s): 1382 -1391 vol.3

[3] Doeringer, W.; Karjoth, G.; Nassehi, M., "Routing on longest-matching prefixes", Networking, IEEE/ACM Transactions on, Volume: 4 Issue: 1, Feb. 1996 Page(s): 86 -97

[4] M. Degermark, A. Brodnik, S. Carlsson, and S. Pink, "Small forwarding tables for fast routing lookups", in Proc. ACM SIGCOMM, Cannes, France, 1997, pp. 3–14.

[5] M. Waldvogel, G. Varghese, J. Turner, and B. Plattner, "Scalable high speed IP routing table lookups", in Proc. ACM SIGCOMM '97, Sept. 1997, pp. 25–3.

[6] V. Srinivasan and G. Varghese, "Faster IP lookups using controlled prefix expansion", in Proc. SIGMETRICS, Madison, WI, 1998, pp.1–10.

[7] Gupta, P.; Lin, S.; McKeown, N., "Routing lookups in hardware at memory access speeds", INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Volume: 3, 29 March-2 April 1998 Page(s): 1240 -1247 vol.3

[8] Lampson, B.; Srinivasan, V.; Varghese, G., "IP lookups using multiway and multicolumn search", INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Volume: 3, 29 March-2 April 1998 Page(s): 1248 -1256 vol.3

[9] H. H.-Y.Henry Hong-Yi Tzeng and T.Tony Przygienda, "On fast address-lookup algorithms", IEEE J. Select. Areas Commun., vol. 17, pp. 1067–1082, June 1999.

[10] Kobayashi, M.; Murase, T.; Kuriyama, A., "A longest prefix match search engine for multi-gigabit IP processing", Communications, 2000. ICC 2000. 2000 IEEE International Conference on, Volume: 3, 18-22 June 2000 Page(s): 1360 -1364 vol.3

[11] N. Yazdani and P. S. Min, "Fast and scalable schemes for the IP address lookup problem", in Proc. IEEE Conf. High Performance Switching and Routing, 2000, pp. 83–92.

[12] Huan Liu, "Reducing routing table size using ternary-CAM", Hot Interconnects 9, 2001. , 22-24 Aug. 2001 Page(s): 69 –73

[13] Marcel Waldvogel, George Varghese, Jon Turner, "Scalable high-speed prefix matching", Bernhard Plattner ACM Transactions on Computer Systems (TOCS) (November 2001) Volume: 19 Issue: 4

[14] Sharma, S.; Panigrahy, R., "Sorting and searching using ternary CAMs", High Performance Interconnects, 2002. Proceedings. 10th Symposium on, 21-23 Aug. 2002 Page(s): 101 -106

[15] Desai, M.; Gupta, R.; Karandikar, A.; Saxena, K.; Samant, V.; "Reconfigurable finite-state machine based IP lookup engine for high-speed router", Selected Areas in Communications, IEEE Journal on, Volume: 21 Issue: 4, May 2003 Page(s): 501 -512

[16] Zane, F.; Narlikar, G.; Basu, A., "CoolCAMs: Power-efficient TCAMs for forwarding engines", INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE, Volume: 1, 30 March - 3

April 2003 Page(s): 42 -52
[17] Sarang Dharmapurikar, Praveen Krishnamurthy, David E. Taylor, “Longest prefix matching using bloom filters”, Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications (August 2003)
[18] Hyesook Lim; Ji-Hyun Seo; Yeo-Jin Jung, “High speed IP address lookup architecture using hashing”, Communications Letters, IEEE, Volume: 7 Issue: 10, Oct. 2003 Page(s): 502 –504
[19] Sangireddy, R.; Somani, A.K., “High-speed IP routing with binary decision diagrams based hardware address lookup engine”, Selected Areas in Communications, IEEE Journal on, Volume: 21 Issue: 4, May 2003 Page(s): 513 –521