IC Design of IPv6 Routing Lookup for High Speed Networks

Yuan-Sun Chu, Hui-Kai Su, Po-Feng Lin, and Ming-Jen Chen

Department of Electrical Engineering,
National Chung-Cheng University, Chia-Yi, Taiwan 621, R.O.C
{chu, pat}@ee.ccu.edu.tw

Abstract. In recent years, there are many researches for routing lookup. Most of them can achieve high average search throughput for IPv4, but they are slow in the updating speed and cannot suit to 128 bits IPv6 address even in hardware architecture. This paper proposed a routing lookup system which contains an ASIC of routing lookup table and off-chip memory sets. In the performance analysis, 91.89 % routing entries of the routing table can be searched in one memory access, and the worst case about 10 % needs two memory accesses. The routing lookup system approaches 213.4 Mlps (109.26 Gb/s). It is enough to satisfy the high speed link OC-768 (40 Gb/s) with 150000 routing entries.

1 Introduction

In recent years, there are many researches for routing lookup. [1, 2] create routing lookup table with trie. [3] proposed hierarchical hardware architecture for IPv4 routing lookup. They can achieve high average search throughput for IPv4, but they cannot suit to 128 bits IPv6 address. [4, 5, 6] proposed routing lookup with CAM, and all match action only needs one clock cycle. But it needs special mechanisms to solve the sorting problem and expensive, especially TCAM.

This paper proposes a routing lookup system which contains an ASIC and off-chip RAM for IPv6. The routing scheme is based on the prefix length distribution of 6Net routing tables. In the proposed system, the lookup speed with 281.69 Mlps can satisfy the requirement of OC-768, and only needs 20.04 KB TCAM, 10.24 KB BCAM, and 29.29 MB RAM for 150000 routing entries.

2 System Architecture

Figure 1 shows the system architecture which composed of an ASIC and a memory set stores routing table. The off-chip RAM is with two hierarchical levels. First level is composed of 3 hash tables, and second level is a pure SRAM table. The hash tables store routing entries with prefix equal to 32, 48 and 64 respectively.

The ASIC has the complete functions, including inserting, searching, updating and deleting. Moreover, it focus on 1-64 bits of IPv6 address (network ID).

V.N. Alexandrov et al. (Eds.): ICCS 2006, Part IV, LNCS 3994, pp. 184–187, 2006.

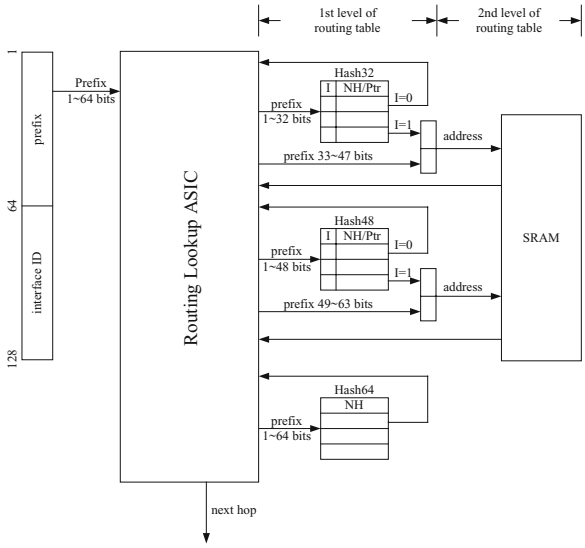

Fig. 1. The routing lookup system

Table 1. The entropy of various hash functions

Hash index	Hash function			
	Bit extraction	Fletch checksum	XOR folding	CRC
12	11.9694	11.9714	11.9753	11.977

80 % hit ratio is guaranteed on the chip. The CAM is used as an on-chip memory for the fast search in the ASIC, and cache replacement algorithm and FIFO are used.

3 Hash Scheme

Table 1 shows the entropy of various hash functions by different schemes. The simulation trace is based on TANET[7] network. It consists of 7.66157 million entries in a period of one hour, and there are 43867 distinct destination addresses. The hash functions are used to simulate Bit Extraction, Fletch Checksum, XOR Folding, and CRC.

Table 1 shows that CRC is the best scheme for the hash function, but it requires complex computation. The XOR folding is also an excellent hash function and simple to be implemented in hardware, so it is used in our scheme.

4 Cache Replacement Algorithm

When a cache misses, the new referenced data needs to be inserted into the cache table. The simulation analyzed five cache replacement algorithms. These

Table 2. The simulation results of cache replacement algorithms

No. of cache entries	LRU	mLRU (4 seg.)	mLRU (16 seg.)	SF-LRU	LFU	FIFO
64	54.72 %	45.84 %	36.20 %	57.15 %	48.65 %	49.21 %
128	61.50 %	49.91 %	42.21 %	64.10 %	55.71 %	56.07 %
512	73.52 %	64.10 %	49.91 %	79.66 %	71.97 %	72.06 %
1024	79.78 %	71.50 %	57.33 %	87.32 %	85.76 %	81.19 %

Table 3. The routing lookup speed in ideal case

Cache hit ratio	Clock period on chip	Clock period off chip	Average routing lookup time	Lookup speed	Provide line rate (64 B)
60 %	3.3 ns	3.3 ns	6.072 ns	164.69 Mlps	84.32 Gb/s
		5 ns	6.82 ns	146.63 Mlps	75.07 Gb/s
		10 ns	9.02 ns	110.86 Mlps	56.76 Gb/s
70 %	3.3 ns	3.3 ns	5.379 ns	185.91 Mlps	95.19 Gb/s
		5 ns	5.94 ns	168.35 Mlps	86.2 Gb/s
		10 ns	7.59 ns	131.75 Mlps	67.46 Gb/s
80 %	3.3 ns	3.3 ns	4.686 ns	213.4 Mlps	109.26 Gb/s
		5 ns	5.06 ns	197.63 Mlps	101.19 Gb/s
		10 ns	6.16 ns	162.34 Mlps	83.12 Gb/s

five algorithms are FIFO [8], LRU [8], mLRU[9], SF-LRU[10], and LFU [11]. The simulation results are shown in Table 2.

In the simulation, SF-LRU and LFU have good performance. But SF-LRU and LFU need counters to record the last reference time and sorting action is too complex in hardware. FIFO has good performance and 81.19 % hit ratio with 1024 entries. It is enough in network traffic. It is simple for hardware design and only needs one register to record the next CAM address.

5 Efficiency Analysis

CCUEE SOC Lab proposes a PF-CDPD CAM [12]. Its clock period can approach 3.3 ns when CAM size is $1024 \times (64 + 8)$. We assume the ideal off-chip clock period can approach the ASIC's speed, i.e., 3.3 ns. We refer to the actual SRAM clock period (5 ns) and CAM clock period (10 ns). The analysis results of lookup speed in ideal case is shown Table 3. The lookup speeds can satisfy the OC-768 requirement. The best ideal lookup speed is 213.4 Mlps with 80 % hit ratio.

Table 4. The comparison with the related works

	Our Scheme	BDD [1]	IPv4/IPv6 dual [5]	Fast TCAM [4]
Implement method	hardware	software	hardware (TCAM)	hardware (TCAM)
Worst search latency	2 memory access	depend on trie depth	7-stage pipeline	3 clock cycle latency, 1 clock is 5 ns
Lookup speed	213.4 Mlps (150000 prefixes)	168.6 Mlps for 29487 prefixes	100 Mlps	200 Mlps
Memory size	TCAM: 20.04 KB, CAM: 10.24 KB, RAM: 29.29 MB	non-available	non-available	21 MB capacity, 21632 entries

6 Conclusion

The IPv6 routing lookup system is proposed with a routing lookup ASIC and a memory set. The scheme is based on the prefix length distribution of 6Net routing tables. The first level in the proposed routing table can cover about 91.89 % routing entries. The ASIC has the complete routing lookup functions: insert, search, update, and delete. A FIFO is used as cache replacement algorithm in the proposed architecture by using a CAM with 1024 entries. It can guarantee 80 % hit ratio, so the speed can approach 213.4 Mlps and satisfy the requirement of OC-768. The system only needs 20.04 KB TCAM, 10.24 KB BCAM, and 29.29 MB RAM for 150000 routing entries. Table 4 shows the comparison with the related works.

References

1. Sangireddy, R., Somani, A.: High-speed IP routing with binary decision diagrams based hardware address lookup engine. IEEE Journal on Selected Areas in Communications **21** (2003) 513 – 521
2. Lampson, B., Srinivasan, V., Varghese, G.: IP lookups using multiway and multicolumn search. IEEE/ACM Transactions on Networking **7** (1999) 324 – 334
3. Huang, N.F., Zhao, S.M.: A novel IP-routing lookup scheme and hardware architecture for multigigabit switching routers. IEEE Journal on Selected Areas in Communications **17** (1999) 1093 – 1104
4. Gamache, B., Pfeffer, Z., Khatri, S.: A fast ternary CAM design for IP networking applications. In: ICCCN 2003. (2003)
5. Wang, Z.X., Wang, H.M., Sun, Y.M.: High-performance IPv4/IPv6 dual-stack routing lookup. In: 18th International Conference on Advanced Information Networking and Applications. Volume 1. (2004)
6. Hayashi, T., Miyazaki, T.: High-speed table lookup engine for IPv6 longest prefix match. In: IEEE Global Telecommunications Conference (GLOBECOM 1999). Volume 2. (1999)
7. The Computer Center of the Ministry of Education in Taiwan: Taiwan academic network (TANET). http://www.edu.tw/EDU_WEB/EDU_MGT/MOECC/EDU0688001/tanet/1.htm (in Chinese) (2005)
8. Tanenbaum, A.S., Woodhull, A.S.: Operating Systems: Design And Implementation, Second Edition. Prentice Hall (1996)
9. Liu, H.: Reducing cache miss ratio for routing prefix cache. In: IEEE Global Telecommunications Conference (GLOBECOM 2002). Volume 3. (2002)
10. Alghazo, J., Akaaboune, A., Botros, N.: SF-LRU cache replacement algorithm. In: Records of the International Workshop on Memory Technology, Design and Testing (MTDT'04). (2004)
11. Shyu, W.L., Wu, C.S., Hou, T.C.: Efficiency analyses on routing cache replacement algorithms. In: IEEE International Conference on Communications (ICC 2002). Volume 4. (2002)
12. Wang, J.S., Li, H.Y., Chen, C.C., Yeh, C.: An AND-type match-line scheme for energy-efficient content addressable memories. In: IEEE International Solid-State Circuits Conference (ISSCC). (2005)