Optimizing Memory Hierarchy within an Internet Router for High-Throughput and Energy-Efficient Packet Processing

Kyosuke Tanaka (Advisors: Hayato Yamaki, Shinobu Miwa, Hiroki Honda)
The University of Electro-Communications
{kyontan,yamaki,miwa,honda}@hpc.is.uec.ac.jp

1 INTRODUCTION

The rapid growth in internet traffic increases the demand for fast and energy-efficient processing in internet routers. Table lookup is known as a bottleneck of improving throughput and energy-efficiency in routers. Typical routers use ternary content addressable memory (TCAM) for fast table lookup; however, it consumes significant amount of energy. Moreover, it is difficult for routers with TCAM to achieve high-throughput over 400 Gbps.

To solve this problem, packet processing cache (PPC), which reuses the results of TCAM accesses, has been proposed [1]. PPC can achieve 448.9 Gbps with 17.9% of energy of conventional architecture by reducing TCAM access count; however, we have more chance to optimize it. For example, multi-level caches and DRAM main memory are widely used in microprocessors, so that such a memory configuration may be applicable to routers. Optimizing memory hierarchy within a router has a great opportunity to improve throughput and energy-efficiency in packet processing, but there is no study that focuses on it. As the first step toward optimal memory configuration, in this paper we introduce multi-level PPC into routers.

2 PPC ARCHITECTURE

A packet arrived at a router is processed with some tables, which are typically configured with TCAM. The five tuple of the packet, which consists of source and destination IP addresses, source and destination port numbers, and protocol number, is used as a key to lookup the tables. The typical data size is 15 bytes contains set of results of TCAM accesses. As five tuple size is 13 bytes, each cache line size is 28 bytes.

Fig. 1 shows the overview of PPC; based on L1 caches in microprocessors, a typical PPC is designed as a 4-way set associative LRU cache that consists of 1K entries. However, the miss rate of conventional PPC (about 20%) is higher than those of L1 caches in microprocessors.

3 EXPERIMENTAL METHODOLOGY

We evaluated various 2-level PPC (L2 PPC) configurations in terms of PPC hit rate, throughput and energy by using our in-house PPC simulator. We used various real-network traces obtained from RIPE Network Coordination Centre [2], mostly 90-second traces of under 1Gbps link between 2004 and 2006, and WIDE MAWI Working Group Traffic Archive [3], 15-minute traces of WIDE trans-pacific 1Gbps link in Japan in 2017 .

4 EXPERIMENTAL RESULT

We tested two write policies (i.e., write-through and write-back) and two hierarchical organizations (i.e., inclusive and exclusive), but present the experimental results of write-through and inclusive PPCs in this paper because we were not able to see any remarkable differences between these parameters. Fig. 2 shows the breakdown of memory accesses for various traces and L2 size. We configured 1K entries for L1 PPC, which was also used in [1]. The figure shows that an L2 PPC with 64K entries can reduce TCAM accesses by 79% on average. Also, we can see that an L2 PPC with 128K+ entries has little advantage when compared to 64K entries.

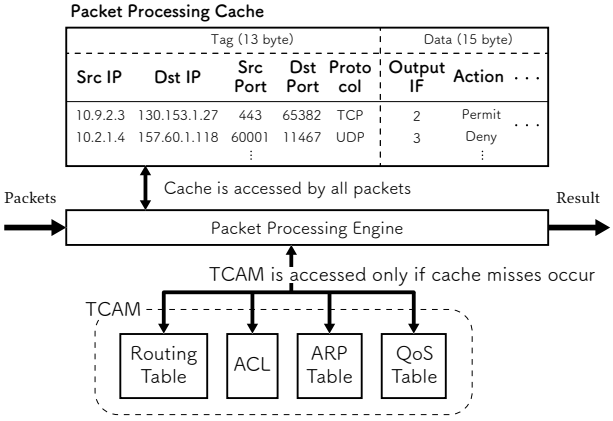

Figure 1: PPC architecture

Next, we show throughput and energy of routers with 2-level PPC. We use L2 PPCs with 8K, 16K and 64K entries for the experiment, while we use a 1K-entry L1 PPC.

Throughput of routers with 2-level PPC (Th_{PPC}) can be calculated as:

$$Th_{PPC} = \frac{512bit}{t_{Avg.}} = \frac{512bit}{\max\{t_{L1}, t_{L2} \cdot m_{L1}, t_{TCAM} \cdot m_{L1} \cdot m_{L2}\}} \quad (1)$$

Here, t_{L1}, t_{L2} and t_{TCAM} represent the latencies of L1 and L2 PPCs and TCAM, respectively; $t_{Avg.}$ represents average lookup latency per packet; and m_{L1}, m_{L2} represent the cache miss rate of L1 and L2 PPCs, respectively. 512-bit means the shortest packet length.

Meanwhile, dynamic energy of table lookup per packet (DE_{PPC}) and average power consumption of table lookup (P_{PPC}) can be calculated as:

$$DE_{PPC} = DE_{L1} + DE_{L2} \cdot m_{L1} + 4DE_{TCAM} \cdot m_{L1} \cdot m_{L2} \quad (2)$$

$$P_{PPC} = DE_{PPC} \cdot n + SP_{L1} + SP_{L2} + SP_{TCAM} \quad (3)$$

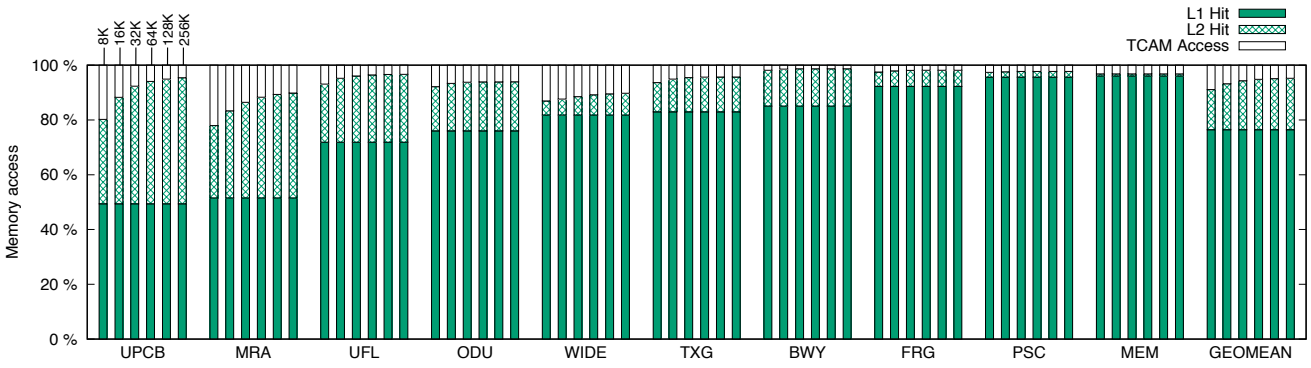

Figure 2: Breakdown of memory accesses.

Here, DE_{L1}, DE_{L2} and DE_{TCAM} represent the dynamic energy per access of L1 and L2 PPCs and TCAM, respectively; SP_{L1}, SP_{L2} and SP_{TCAM} represent their static power, respectively. n represents the number of packets arriving at the router per second. We assumed that a packet needs to access four tables as shown in Fig. 1, thus we multiplied the dynamic energy of TCAM by 4.

Latency, access energy and static power of PPCs are estimated using CACTI 6.5 [4] configured with each one read and write exclusive port, 32nm technology and itrs-hp profile for L1 cache and itrs-lstp profile for L2 cache. We used the conventional TCAM power and delay models shown in [5] as t_{TCAM}, DE_{TCAM} and SP_{TCAM}, note that we consider TCAM has 512K entries by multiplying dynamic energy and static power.

Table 1: Throughput and dynamic energy of PPCs

	L1: 1K L2: 8K	L1: 1K L2: 16K	L1: 1K L2: 64K	Conventional (L1: 1K)
Th_{PPC} [Gbps]	571.2	627.4	393.5	215.6
DE_{PPC} [nJ]	0.989	0.814	0.778	2.501

Tab. 1 shows our estimation results. As shown in the table, introducing L2 PPC improves both throughput and dynamic energy efficiency. However, an L2 PPC with 64K entries is not the best configuration from the aspect of the throughput though it achieves the lowest PPC miss rate. This indicates that introducing L2 PPC moves the bottleneck of throughput to L2 PPC accesses. The PPC configured as 1K-entry L1 cache and 16K-entry L2 cache improves throughput and dynamic energy of table lookups by 191.0% and 67.5%, respectively, when compared to conventional PPC.

Next, we analyzesd the breakdown of power consumptions in PPC under 400Gbps traffic and showed the result in Fig. 3. The 400Gbps traffic was created based on the 1Gbps WIDE traffic by multiplying the number of packets. Here, DP_{L1}, DP_{L2} and DP_{TCAM} represent the power consumed by the dynamic energy of L1 and L2 PPCs and TCAM, respectively. Fig. 3 indicates that the dynamic power of TCAM is dominant in PPC and the static powers of memories are relatively small. Note that the power of L2 PPC is smaller than L1 PPC because of difference of memory profiles between

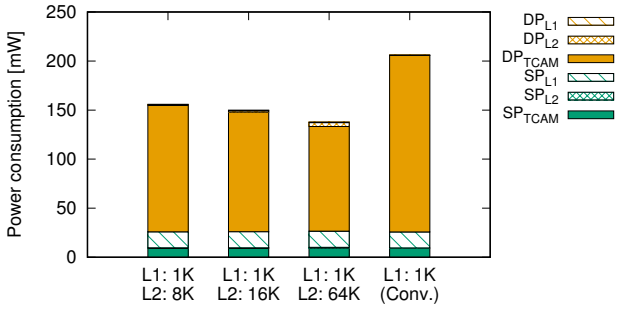

Figure 3: Breakdown of memory power consumptions.

L1 and L2 PPCs. From this result, increasing the L2 cache size is effective to reduce the power consumption of PPC.

In summary, the PPC configured as 1K-entry L1 cache and 16K-entry L2 cache improves both throughput and energy efficiency of table lookups by 191.0% and 27.3%, respectively, when compared to conventional PPC.

5 CONCLUSIONS

This paper reported that multi-level PPC can improve throughput and energy-efficiency in internet routers. Our experimental results showed that introducing L2 PPC into routers can reduce cache misses by about 80%. In addition, we tested various L2 PPC and found that multi-level PPC can improve throughput and energy-efficiency of the table lookups in routers by up to 191.0% and 27.3%, respectively.

REFERENCES

[1] H. Yamaki, et al., Data prediction for response flows in packet processing cache, *DAC*, pp.1-6, 2018.
[2] RIPE Network Coordination Centre, Réseaux IP Européens Network Coordination Centre, Available: https://www.ripe.net/. [Accessed Jul. 22, 2018]
[3] WIDE MAWI WorkingGroup 'MAWI Working Group Traffic Archive' Available: http://mawi.wide.ad.jp/mawi/. [Accessed Jul. 22, 2018]
[4] N. Muralimanohar et al., Optimizing NUCA organizations and wiring alternatives for large caches with CACTI 6.0, *Proc. of MICRO 40*, pp.3-14, 2007.
[5] Y. Cheng et al., Low leakage mask vertical control TCAM for network router, *Proc. of IEEE APCCAS* , pp. 469-472, 2016.