

# 1Tbps 実現に向けたルータのメモリ階層の最適化

田中 京介<sup>1,a)</sup> 八巻 隼人<sup>1</sup> 三輪 忍<sup>1</sup> 本多 弘樹<sup>1</sup>

**概要:** インターネットトラフィックは年々増加しており, 近い将来には 400Gbps や 1Tbps といった超大容量回線が登場することが見込まれている. しかしながら, 既存のインターネットルータの packets 処理スループットは 100Gbps 程度が限界であり, 1Tbps 達成のためには大幅なスループットの向上が必要とされる. これに対し, インターネットルータの packets 処理においてスループット上のボトルネックとなっているテーブル検索を, キャッシュを用いて高速化する packets 処理キャッシュ (PPC: Packet Processing Engine) が提案されている. PPC を用いたテーブル検索処理では, PPC のキャッシュミス率がスループットを決定するため, これまでキャッシュミスを削減する様々な研究が行われてきた. 一方で, キャッシュそのものの構造に焦点を当てた研究は行われておらず, キャッシュの最適な階層や容量については議論されていなかった. 本報告では 2 階層と 3 階層の PPC についてそれぞれ実ネットワークトレースを用いたシミュレーションを行い, 複数階層を持つ PPC の最適な構成について検討した. 評価の結果, 3 階層を持つ PPC の最適な構成ではテーブル検索処理のスループットを 254.1% 改善し, 1Tbps を達成できることが分かった. また, ネットワークトレースによっては大容量のラストレベルキャッシュを導入しても初期参照ミスが大きいことにより 1Tbps を達成できないものがあり, さらなる性能改善には初期参照ミスの削減が不可欠であることを示した.

**キーワード:** ネットワークルータ, packets 処理, キャッシュ, メモリ階層

## 1. はじめに

近年, インターネットトラフィック量は急激に増加しており [1], 将来的なインターネットルータの処理スループット不足が懸念されている. 現在のハイエンドルータの処理能力はポートあたり 100Gbps 程度であり, 数年内に標準化が完了する 400 Gbps Ethernet 規格 [2] や, 実現に向けたプロジェクトが進んでいる 1Tbps 回線に対応するためには, 大幅に処理スループットを向上しなければならない. 更には, 処理 packets 数の増加に伴いインターネットルータの消費電力が増大することも問題視されており, 次世代ルータには 1Tbps 超の packets 処理を低電力に実現できることが求められている.

インターネットルータでは, テーブル検索処理が packets 処理スループット上のボトルネックとなっている [3]. インターネットルータの packets 処理におけるテーブル検索処理とは, ルータへ到着した packets の転送に要する情報 (宛先, フィルタリング, IP 変換等) を, ルータ内のテーブルを検索することで得る処理を指す. 一般的なイン

ターネットルータのテーブル検索処理では, ルーティングテーブル, ARP (Address Resolution Protocol) テーブル, ACL (Access Control List), そして QoS (Quality of Service) テーブルといった複数テーブルの検索を要する.

近年のハイエンドルータは, これらのテーブルを TCAM (Ternary Content Addressable Memory) と呼ばれる 1 サイクルで検索処理を行うことができる高速なメモリへ格納することにより, 高いテーブル検索性能を実現している. しかしながら, 現在の TCAM の検索性能では, 最小サイズの packets が連続してルータへ到着するワーストケースにおいて 100Gbps 程度のスループットの達成が限界であり, TCAM であっても将来的な回線速度に対する検索性能の不足が懸念される. 加えて, TCAM は一般的に使用されるメモリである SRAM (Static Random Access Memory) に比べ, 1 回のアクセスに必要な消費電力が極めて大きく, ルータ全消費電力の 40% 程度を TCAM が占めていると報告されている [4][5]. これらのことから, インターネットルータが 1Tbps のスループットを省電力に達成するためには, 既存の TCAM の性能改善だけでなく, TCAM に依らない新たなテーブル検索アーキテクチャが必要となる.

テーブル検索処理を高速化する手法として, 小容量かつ高速な SRAM をキャッシュメモリとして用いるパケッ

<sup>1</sup> 電気通信大学  
1-5-1, Chofu Ga Oka, Chofu-shi, Tokyo 182-8585, Japan  
<sup>a)</sup> kyontan@hpc.is.uec.ac.jp

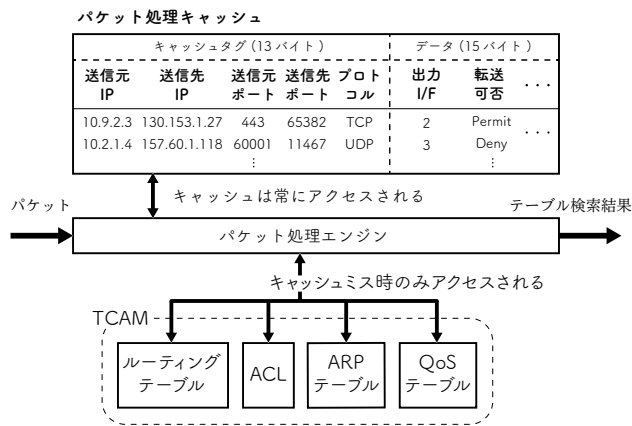


図 1: パケット処理キャッシュを用いたテーブル検索処理の流れ

ト処理キャッシュ (Packet Processing Cache: PPC) が提案されている。PPC は、TCAM 検索結果をキャッシュに記憶して再利用することで、TCAM のアクセス回数を削減し、高速かつ低消費電力なテーブル検索処理を実現する。ネットワーク上のトラフィックには時間的局所性があり、同一の TCAM エントリを参照するパケットが連続してルータへ到着する傾向にあることから、PPC を用いることで TCAM アクセス回数を大幅に削減できる可能性がある。これまで PPC のキャッシュミス削減する様々な方法が研究されてきた。しかしながら、既存研究により達成されるスループットは 400Gbps 程度であり、いまだ 1Tbps を達成できるような PPC の構造は検討されていない。そこで本研究では、PPC のキャッシュ構造を階層化することで、1Tbps を実現しうる PPC の構造について検討し、評価する。

## 2. パケット処理キャッシュ (PPC)

テーブル検索処理では、パケットヘッダ内の 5 タプル (送信元 IP アドレス, 送信先 IP アドレス, 送信元ポート, 送信先ポート, プロトコル番号) の一部分あるいは全てをキーとして用いて各テーブルの検索を行う。PPC は、上記の 5 タプルによりパケットをフローと呼ぶ単位で分類し、フローによって決定される全テーブルの検索結果を 1 エントリに格納する。これにより、フローの 1 パケット目のテーブル検索処理の結果をキャッシュへ格納し、同一フローの 2 パケット目以降のテーブル検索結果を PPC から得ることが可能となる。キャッシュメモリとして利用される SRAM は TCAM に比べてアクセス時のレイテンシと消費エネルギーが少ないことから、PPC を用いることでテーブル検索処理の高スループット化と省電力化が期待できる。

図 1 に PPC を用いたテーブル検索処理の概要を示す。PPC では、フローを決定する 13 バイトの 5 タプル情報をキャッシュタグとして用い、各テーブルの検索結果をデー

タとして保持する。図に示した 4 つのテーブルを持つ一般的な構成のインターネットルータでは、出力先ポートとして 1 バイト、送信元・宛先 MAC アドレスとして 12 バイト、ACL によるフィルタリング結果として 1 バイト、QoS 情報として 1 バイトの計 15 バイトをデータとして保持しており、4 回の TCAM アクセスを 1 回のキャッシュアクセスにより代替している。

PPC を用いた際のテーブル検索性能は、PPC のレイテンシとキャッシュミス率により決定される。一般的に、キャッシュ容量を大きくすることでキャッシュミス率は改善可能だが、容量増加に伴いレイテンシも増加する。そのため、従来の PPC は 1K エントリ、すなわち 28KB 程度の小容量なキャッシュを用いており、小容量なキャッシュでいかに低いキャッシュミス率を達成するかが重要であった。

## 3. 関連研究

PPC におけるキャッシュミスは初期参照ミス、容量性ミス、そして衝突性ミスに分類され、それぞれに対して様々な改善手法が提案されてきた。

容量性ミスに着目した研究として、Cheng らはキャッシュタグを圧縮することによるエントリあたりのサイズの削減手法を検討している [6]。ルータ内の各テーブルで用いられるフロー情報は 104bit と大きいことから、32bit のハッシュ値をタグ情報として用いる Digest Cache を提案している。Digest Cache はブルームフィルタをルータで用いるために改変した手法であり、許容される割合の誤ったテーブル検索結果と引き換えに高い容量あたりのエントリ数を実現している。

衝突性ミスに焦点を当てた研究として、Kim らは一般的に使用されるエントリ置換アルゴリズムである LRU は時間的局所性のみを考慮しており、インターネットルータで考慮すべきネットワークの特性を反映しておらず高い効果を発揮できないと述べている [7]。そこで、少なくとも 1 回は参照されたエントリを Switching Entry, 1 回も参照されていないエントリを Non Switching Entry として区別し、Non Switching Entry を優先的に追い出す手法を提案している。初めに提案している Weighted Priority LRU Scheme では、Non Switching Entry の中でも挿入されてから時間が経っていないエントリを追い出しの対象外とすることで、参照される可能性の高い Non Switching Entry を優先的に残している。また、L2A scheme として過去 2 回のパケットのタイムスタンプの合算値が小さいエントリを追い出す手法を提案している。これらについて、キャッシュ中に占める Switching Entry の数およびキャッシュの平均置換回数の観点から評価を行っており、特に L2A scheme は LRU と比べてより長期間 Switching Entry を保持し、キャッシュの平均置換回数が優れているとした。しかしながら、タイムスタンプを保持するためのメモリや

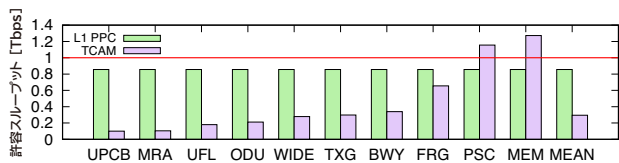


図 2: 1 階層 PPC における PPC と TCAM の許容スループット

ハードウェアのコストには触れられておらず、メモリ容量の少ない PPC ではメモリコストの増大が致命的な問題になると考えられる。

我々も、過去に PPC のミス削減手法として、フローの構成パケット数に着目した新たなエントリ置換方式の検討 [8]、初期参照ミスの削減を目的としたキャッシュエントリの挿入方式の検討を行った [9]。論文 [8] では、パケット数の多いフローのエントリを長期間保持し、少数のパケットからなるフローのエントリを迅速に追い出すエントリ置換アルゴリズムとして ELC (Elevator Cache) を提案した。これにより、PPC のミス率を LRU と比べて平均 11.1%、削減している。また論文 [9] では、初期参照ミスの削減のため、トラフィックにサーバ・クライアント型の通信が多いことに着目した投機的なキャッシュの挿入方式を提案している。実ネットワークにおけるトラフィックでは、およそ 70% が対称なフロー、つまり送信元と送信先の IP アドレスとポートを入れ替えたフローから構成されている。この特性を利用し、フローの最初のパケットを処理するとき、その後に対称なフローのパケットが来ることを期待して対称なフローのエントリを投機的にキャッシュする RPC (Response Prediction Cache) を提案し、キャッシュミスを既存の PPC に比べ平均 15.3% 削減している。

#### 4. PPC の多階層化

PPC と同様にキャッシュ機構を有する CPU ではキャッシュを多層化することで、前段の小容量なキャッシュにより高速化の恩恵を受けつつ、後段の大容量なキャッシュにより容量性ミスを削減し、キャッシュ全体としてのミス率を下げるアーキテクチャが一般的となっている。しかしながら、PPC においてはキャッシュの階層化ははまだ検討されておらず、これらの手法をルータのパケット処理に適用することで、テーブル検索処理を高速化できる余地がある。そこで、本研究では PPC における最適なキャッシュ階層を検討することで、ルータの更なるスループット向上の可能性を模索する。

まず、我々は 1 階層の PPC におけるスループットについて分析を行い、1Tbps を達成する上でのボトルネックを明らかにした。図 2 は、後述する PPC シミュレータを用いて 10 種の実ネットワークトレースに 1 階層の PPC を適用した場合の、PPC と TCAM の許容スループットを示して

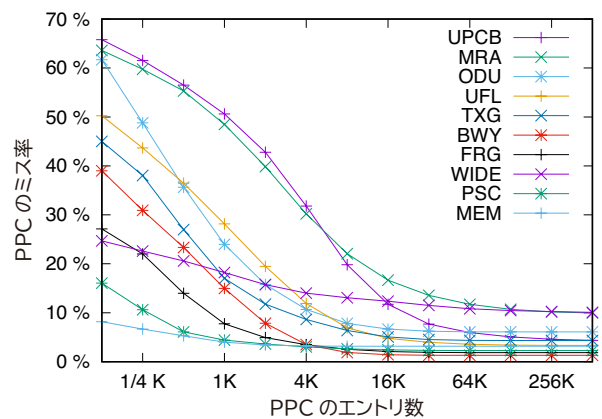


図 3: PPC のエントリ数によるミス率の変化

いる。許容スループットとは、その階層のメモリがパケットをロスすることなく処理可能なルータの入力トラフィック容量である。また、ルータのスループットは、最も許容スループットの低いメモリによって律速される。図 2 より、まず、1 階層の PPC はキャッシュのレイテンシが大きく、いかにキャッシュミスを削減しても 1Tbps を達成できないことがわかる。これはすなわち、より小容量で低レイテンシなキャッシュがファーストレベルキャッシュとして必要となることを意味する。更に、多くのトレースにおいて TCAM の許容スループットが不十分であることがわかる。これは、1 階層の PPC では十分なキャッシュヒットが得られておらず、TCAM アクセスがスループット上のボトルネックとなっていることを示している。従って、大容量なラストレベルキャッシュを用いてキャッシュミスを削減し、TCAM の許容スループットを 1Tbps まで向上させる必要がある。

次に、1 階層の PPC のエントリ数によるミス率の変化を測定し、図 3 に示した。図に示されているように、PPC におけるミス率の傾向はネットワークにより大きく異なる一方で、128K エントリ以上の PPC ではどのトレースでもほぼミス率の改善は見られない。これは、PPC が最大でも 128K エントリ程度を持てば十分であることを示しており、これ以上のミス削減には初期参照ミスの削減が必要であることを示している。

#### 5. 多階層 PPC の評価

##### 5.1 評価手法

本報告では、多階層の PPC のメモリ構成を最適化することにより、最も少ない消費電力で 1Tbps のスループットを達成する PPC のメモリ構成を探索した。

我々は、多階層の PPC のミス率を評価するため、PPC シミュレータをソフトウェアにより実装した。このシミュレータは PPC のキャッシュ構成およびネットワークトレースを指定することで、そのネットワークトレースで PPC

表 1: 評価に用いた CACTI 6.5 のパラメータ

設定項目	設定値
半導体プロセス	32nm
ポート	1 (read) + 1 (write)
トランジスタモデル (LLC 以外)	ITRS-HP
(LLC)	ITRS-LSTP

を用いてテーブル検索を行った際の、各階層メモリのヒット率と TCAM アクセス率を計測することができる。本報告では PPC 階層間のエントリ制御にはライトスルー方式およびインクルーシブキャッシュを採用した。

PPC のレイテンシや消費電力の見積もりには CACTI 6.5[10] を用いた。評価に用いた CACTI のパラメータを表 1 に示す。ラストレベルキャッシュ (LLC) に関しては、プロセッサの LLC と同様に低消費電力性を重視し、ITRS-LSTP トランジスタモデルを採用している。LLC 以外に関してはアクセス速度を重視し、ITRS-HP トランジスタモデルを採用した。

多階層化した PPC におけるスループットのモデル式として、2 階層の場合のスループット  $Th_{PPC}[Gbps]$  を式 (1) に示した。以下では簡単のため、PPC の  $n$  階層目を  $Ln$  PPC と称する。

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

ここで  $t_{L1}$ ,  $t_{L2}$ ,  $t_{TCAM}$  はそれぞれ L1 PPC, L2 PPC, TCAM のレイテンシを表している。また、 $m_{L1}$  と  $m_{L2}$  は L1 PPC および L2 PPC のミス率を表している。 $l$  は 1 パケットのサイズを表しており、本報告では最悪ケースのシナリオを想定し、最小パケット長である 512bit とした。

同様に、2 階層の PPC におけるパケットあたりの消費エネルギー  $DE_{PPC}$  および平均消費電力  $P_{PPC}$  のモデル式は以下となる。

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

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

ここで  $DE_{L1}$ ,  $DE_{L2}$ ,  $DE_{TCAM}$  はそれぞれ L1 PPC, L2 PPC, TCAM の動的消費エネルギーを表している。また、 $SP_{L1}$ ,  $SP_{L2}$ ,  $SP_{TCAM}$  はそれぞれ L1 PPC, L2 PPC, TCAM の静的消費電力を表している。 $a$  は 1 パケットのテーブル検索処理に要する TCAM アクセス数で、本報告では図 1 に示した一般的なインターネットルータを想定し、 $a = 4$  とした。 $n$  は 1 秒あたりの平均パケット数を表している。なお、TCAM のレイテンシと動的消費エネルギー、静的消費電力の値については論文 [11] を参照した。TCAM についてはルータで一般的に用いられている 20Mbit の容量を持つものと仮定し、また、TCAM の動的消費エネルギーおよび静的消費電力は容量に比例するものと仮定して

表 2: 探索したキャッシュエントリ数の範囲

	L1	L2	L3
1 階層 PPC	1/4 K - 512K	N/A	N/A
2 階層 PPC	1/16 K - 8K	8K - 512K	N/A
3 階層 PPC	1/16 K - 1/2 K	1K - 16K	8K - 512K

文献中の値を定数倍した。また、上式では 2 階層 PPC に関する場合の計算式を提示したが、このモデルは 1 階層、3 階層の PPC の場合にも容易に拡張することができる。

シミュレーションでの評価には、複数の実ネットワークにおけるトレース [12][13] を使用した。[12] は WIDE MAWI Working Group Traffic Archive から取得したものであり、2017 年に太平洋を横断する 1Gbps のリンクから取得された 15 分間のトレースである。平均消費電力については、1Gbps のリンクである WIDE のネットワークトレースにおける 1 秒あたりの平均パケット数を 1000 倍し、仮想的な 1Tbps のトラフィックとみなして計算を行った。また、その他の 9 種のトレースについては、RIPE Network Coordination Centre[13] から取得したものであり、2004 年から 2006 年に学術用バックボーンネットワークにおいて取得され、そのほとんどが 1Gbps 以下かつ 90 秒のトレースである。

面積に関する評価では、インターネットルータのチップ内において大部分を占めるとされている [14] パケットバッファと PPC の各キャッシュメモリが占める面積について比較した。インターネットルータのパケットバッファの容量は、商用の 100Gbps ルータ [15] を参考とし 16MB とした。

表 2 にシミュレーションにおいて評価した PPC の各レイヤにおけるエントリ数を示す。PPC のエントリ数は、基本的には 1/16~512K の範囲で変化させた。ただし、複数階層の構成においては、下位のキャッシュの容量が上位のキャッシュ容量よりも大きくなる構成のみを評価した。

## 5.2 最適なエントリ数の検討

我々は多階層の PPC について最適なエントリ数を検討するため、まず 2 階層の PPC を検討した。図 4 に、2 階層の PPC における各構成ごとにスループットと消費電力をプロットしたものを示す。最適な結果として、L1 PPC が 1K エントリかつ L2 PPC が 16K エントリするとき、スループットと平均消費電力はそれぞれ 856.1Gbps, 335.9mW となり、既存の PPC に比べスループットは 190.4% 向上し、平均消費電力は 27.3% 削減された。興味深いことに、最適な構成においても L1 PPC のエントリ数は既存の PPC と同じ 1K エントリとなった。これは L1 PPC のエントリ数を小さくすると L2 PPC のアクセス頻度が増え、それにより全体の許容スループットが最適化されなかったことを示唆している。また、L1 PPC が依然として 1Tbps のテ

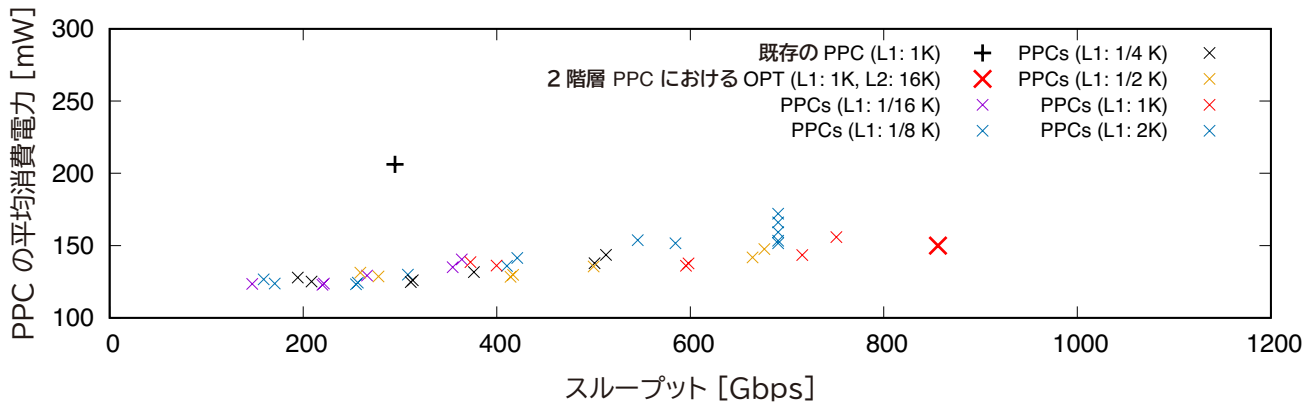


図 4: 2 階層 PPC のそれぞれの構成におけるスループットと平均消費電力

表 3: 3 階層の PPC と既存の PPC の構成, スループット, および, 平均消費電力

	既存の PPC	OPT
L1 エントリ数	1K	1/4 K
L2 エントリ数	N/A	4K
L3 エントリ数	N/A	128K
スループット [Gbps]	294.8	1043.9
平均消費電力 [mW]	476.8	344.2

ブル検索処理を実現する検索性能を持っていないことを示しており, PPC を 3 階層化することでそれぞれの PPC における許容スループットを最適化する必要があることが分かった.

3 階層の PPC について, 図 5 に各構成ごとにスループットと消費電力をプロットしたものを示す. 特に, L1 PPC, L2 PPC, および L3 PPC のエントリ数がそれぞれ 1/4 K, 4K, 128K のとき, スループットの全トレース平均は 1043.9Gbps となる. これは 1Tbps を達成する最も消費電力が少ない構成である. また, 平均消費電力は 182.7mW となり, これは既存の PPC に比べ 11.4% 少なく, 1Gbps あたりの平均消費電力は 75.0% 改善した.

図 6 に, この構成における各レイヤごとの許容スループットを示す. また, 表 3 に最適な 3 階層の PPC と既存の PPC の構成, スループット, および, 平均消費電力の違いをまとめる. 図 6 より, 半分のネットワークトレースでは 1Tbps を達成できておらず, そのうち 3 つ (MRA, WIDE, ODU) では TCAM の許容スループットが 1Tbps に満たないことが分かった. 図 3 に示したように, これら 3 つのトレースに関しては, 128K エントリの PPC を導入した場合でも依然として TCAM へのアクセス頻度が高い. すなわち, これら 3 つのトレースにおける PPC ミスの多くが初期参照ミスであり, これらのトレースに対して 1Tbps を実現するためには初期参照ミスの削減が必須であることを表している.

また, 1Tbps を達成できていないトレースの多くは L2

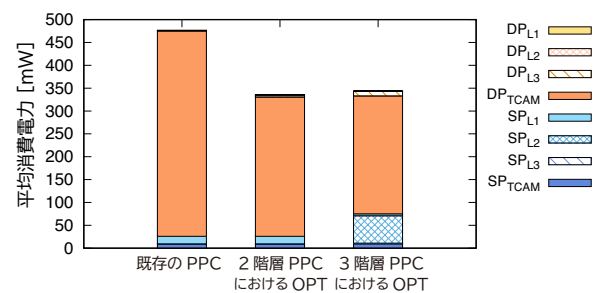


図 8: 各 PPC の構成における消費電力の内訳

表 4: PPC の面積に関する比較

	既存の PPC	OPT	OPT <sub>Area</sub>
L1 PPC [mm <sup>2</sup> ]	0.22	0.14	0.14
L2 PPC [mm <sup>2</sup> ]	N/A	0.73	0.73
L3 PPC [mm <sup>2</sup> ]	N/A	19.04	5.41
合計 [mm <sup>2</sup> ]	0.22	19.90	6.27
パケットバッファ [mm <sup>2</sup> ]		29.47	

PPC の許容スループットが不足している. これは L1 PPC のヒット率が不足していることを示しており, エントリ数の少ない L1 PPC において有用なフローのエントリを長期間保持し, 逆に不要となったエントリを短時間で追い出す手法の検討が重要となってくることを示している.

### 5.3 消費電力に関する分析

図 8 に PPC の消費電力の内訳を示す. PPC を多階層化することは, 主に TCAM の動的消費エネルギーを削減することに貢献していることが分かる. 2 階層 PPC における L2 PPC と 3 階層 PPC における L3 PPC による消費電力のオーバーヘッドはほとんど見られない. これは, LLC に低消費電力のトランジスタモデルを適用したためである.

### 5.4 面積に関する分析

表 4 に示したように, 3 階層 PPC における最適な構

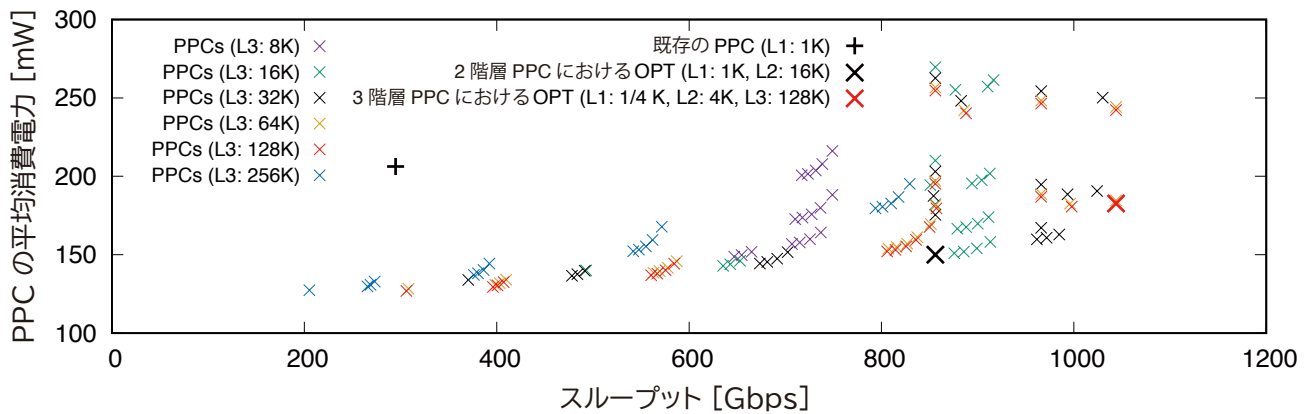


図 5: 3 階層 PPC のそれぞれの構成におけるスループットと平均消費電力

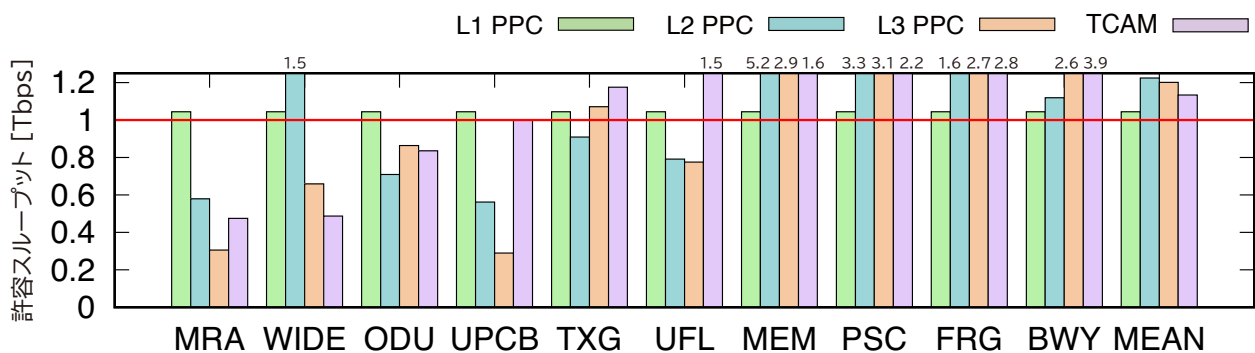


図 6: L3 OPT におけるネットワークトレースごとの許容スループットとボトルネック

成では PPC の回路面積は  $19.90\text{mm}^2$  である。CACTI 6.5 を用いて商用の 100Gbps インターネットルータ [15] のパケットバッファの面積を見積もった結果、16MB のパケットバッファの面積は  $29.47\text{mm}^2$  であった。よって、3 階層 PPC における最適な構成では、パケットバッファのおよそ 67.5% に相当する回路面積が増加する。上記の回路面積の増分の 95.7% が L3 PPC によるものであり、回路面積の観点から言えば、128K エントリの PPC の搭載はやや現実味に欠ける。

一方で、図 3 で示したように、大多数のネットワークトレースでは、32K エントリよりも大きなサイズの PPC を採用することによるキャッシュミス削減効果はほとんどない。そこで、PPC の総面積を  $10\text{mm}^2$  以下とする設計制約を新たに追加し、最適な構成  $\text{OPT}_{\text{Area}}$  を探索したところ、L1 PPC, L2 PPC, および L3 PPC のエントリ数がそれぞれ 1/4 K, 4K, 32K からなる構成が最適となった。このときの PPC の総面積は  $6.28\text{mm}^2$  とパケットバッファの 20% 程度となり、この程度の増分であれば許容できる可能性が高い。また、上記の構成の平均消費電力は  $365.4\text{mW}$  となった。この平均消費電力は前述した最適な 3 階層 PPC の構成に比べると 6.2% の増加ではあるものの、既存の PPC に対しては 23.4% 少ないことから、 $\text{OPT}_{\text{Area}}$  は電力的にも許

容可能と言える。

## 6. 結論

インターネットのトラフィックは年々増加する傾向にあり、次世代のルータでは 1Tbps を超えるパケット処理を低電力に実現できることが求められている。インターネットルータにおいてパケット処理のスループットを向上する上でのボトルネックとして、TCAM によるテーブル検索処理がボトルネックとなっている。TCAM のみによるテーブル検索処理では 100Gbps 程度のスループットの達成が限界であり、これに対処するため小規模なキャッシュメモリを用いたパケット処理キャッシュ (PPC) が提案されている。PPC は TCAM によるテーブル検索処理の結果の一部をキャッシュへ保持することで、時間的局所性を持つネットワークトラフィックにおいて高速なテーブル検索処理を実現している。PPC を用いたルータにおけるスループットおよび消費電力はキャッシュのミス率に依ることから、PPC における様々なキャッシュミス削減手法が研究されてきた。しかしながら、これまでの研究では容量性ミスの本質的な削減に着目した研究は行われてこなかった。また既存の 1 階層の PPC についてボトルネックを検討した結果、PPC のスループットが 1Tbps に満たず、ま



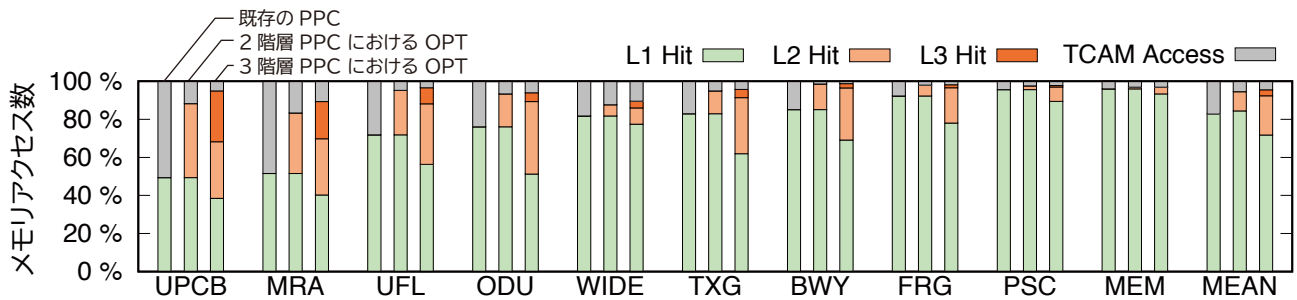


図 7: 各 PPC の構成におけるネットワークトレースごとのメモリアクセスの内訳

たキャッシュミス率も要求に満たないことがわかった。そこで PPC を構成するキャッシュメモリの階層化に着目し、PPC における最適なメモリ階層について検討する。本報告では、PPC シミュレータをソフトウェアで実装し、様々な構成を持つ PPC について検討した。シミュレーションの結果、3 階層の PPC において L1 PPC, L2 PPC, および L3 PPC のエントリ数がそれぞれ 1/4 K, 4K, 128K のとき、1Tbps のスループットを達成できることが分かった。また、このときの平均消費電力は 182.7mW となり、これは既存の PPC に比べ 11.4% 少なく、1Gbps あたりの平均消費電力は 75.0% 改善している。これにより、PPC を多階層化し、適切なエントリ数を選択することによりインターネットルータのテーブル検索処理において 1Tbps のスループットが達成できることがわかった。

謝辞 本研究は、JSPS 科研費 (JP18K18022) による助成を受けたものである。

#### 参考文献

[1] Cisco, “Cisco Visual Networking Index: Forecast and Methodology 2016-2021,” (2017), <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html> (参照 2018-11-3).

[2] IEEE P802.3bs Task Force, “IEEE P802.3bs 400 Gb/s Ethernet Task Force,” (2018), <http://www.ieee802.org/3/bs/> (参照 2018-11-3).

[3] W. Eatherton et al., “Tree Bitmap: Hardware/Software IP Lookups with Incremental Updates,” ACM SIGCOMM Computer Communication Review, 2004, vol. 34, no. 2, pp. 97–122.

[4] M. Nawa et al., “Energy-efficient high-speed search engine using a multi-dimensional TCAM architecture with parallel pipelined subdivided structure,” 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2016, pp. 309–314.

[5] HP Enterprise Group, “Energy Efficient Networking - Business White Paper,” (2011), <http://h17007.www1.hp.com/docs/mark/4AA3-3866ENW.pdf> (参照 2018-11-3).

[6] F. Chang, W. Feng, W. Feng and K. Li, “Efficient Packet Classification with Digest Caches,” Network Processor Design, Volume 3: Issues and Practices (The Morgan

Kaufmann Series in Computer Architecture and Design) 2005, pp.33–54

[7] N.Kim, S.Jean, J.Kim and H.Yoon, “Cache replacement schemes for data-driven label switching networks,” In Proc. of 2001 IEEE Workshop on High Performance Switching and Routing, Dallas, TX, USA, 2001, pp.223–227

[8] H.Yamaki and H.Nishi, “Line Replacement Algorithm for L1-scale Packet Processing Cache,” In Adjunct Proc. of the 13th Int’l Conf. on Mobile and Ubiquitous Systems: Computing Networking and Services (MOBIQUITOUS 2016), Hiroshima, Japan, 2016, pp. 12–17.

[9] H.Yamaki, H.Nishi, S.Miwa and H.Honda, “Data prediction for response flows in packet processing cache,” In Proc. of the 55th Annual Design Automation Conference (DAC ’18), San Francisco, CA, USA, 2018, Article 110, pp.1–6.

[10] N.Muralimanohar, R.Balasubramonian and N.Jouppi, “Optimizing NUCA organizations and wiring alternatives for large caches with CACTI 6.0,” Proc. of the 40th Annual IEEE/ACM Int. Symposium on Microarchitecture (MICRO 40), Chicago, USA, 2007, pp.3–14.

[11] Y. Cheng, J. Chen, T. Wu and Y. Chang, “Low leakage mask vertical control TCAM for network router,” In Proc. of the 2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Jeju, South Korea, 2016, pp.469–472.

[12] WIDE MAWI WorkingGroup, “MAWI Working Group Traffic Archive,” (2018) <http://mawi.wide.ad.jp/mawi/> (参照 2018-11-4).

[13] RIPE Network Coordination Centre, Réseaux IP Européens Network Coordination Centre, “NLNLR AMP Data” ,(2010) <https://labs.ripe.net/datarepository/data-sets/nlanr-amp-data> (参照 2018-11-4).

[14] H. Hassan, M. Said and H. Kim, “Performance evaluation of buffer sharing routers for Network on Chip,” In Proc. of 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS), Abu Dhabi, UAE, 2016, pp.1–4.

[15] Cisco, “Cisco Nexus 9500 R-Series Line Cards and Fabric Modules White Paper,” (2018), <https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-738392.html> (参照 2018-11-4).