

令和 2 年 6 月 8 日現在

機関番号：14401

研究種目：基盤研究(B) (一般)

研究期間：2017～2019

課題番号：17H01733

研究課題名(和文)高性能と低消費電力を両立する情報指向ネットワーク用ルータアーキテクチャ

研究課題名(英文)Architecture for High Performance and Energy Efficient ICN Router

研究代表者

長谷川 亨 (Hasegawa, Toru)

大阪大学・情報科学研究科・教授

研究者番号：70576264

交付決定額(研究期間全体)：(直接経費) 14,200,000円

研究成果の概要(和文)：次世代インターネットアーキテクチャの情報指向ネットワークに対して、100万パケット/秒のパケット転送を実現するルータを実現した。

第一に、低速メモリ装置に蓄積された制御データ構造へのアクセス待ちを隠蔽するプリフェッチアルゴリズムを設計し、プロトタイプ実装することで、汎用のCPUを用いて世界最速の68.3 Gビット/秒の転送速度を実現した。第二に、コンテンツが大容量化してもヒット率を低下させないキャッシュ方式を追加した。第三に、CPUコアに均等に受信パケットを割り当てられるように、複数のCPUコアが同時に同じフロー状態にアクセスしても、一貫性を損なわないデータ構造を設計した。

研究成果の学術的意義や社会的意義

多数の端末を収容し、大量のパケットを転送する今後のインターネットでは、専用LSIを用いた現状のハードウェアルータは、消費電力がトラフィック量に対して増加するため、使い続けることは難しい。本研究では、消費電力の少ない汎用CPUでハードウェアルータと同等のパケット転送速度を実現する転送方式を本研究で開発した。本成果は、高速性の省電力性を両立する次世代ルータの礎となることが期待される。

研究成果の概要(英文)：The study designs and implements a high-speed software router for ICN (Information Centric Networking) as a future Internet architecture. The contributions of the study are three-folded: First, the study designs a pre-prefetch algorithm of data structures for packet forwarding. This enables the packet forwarding speed of 68.3 Gbps, which is the fastest speed for ICN software routers. Second, the study designs a load balancing algorithm for dispatching packets to CPU cores of a multi-core CPU based ICN software router. This enables high speed forwarding for various input traffic patterns. Third, the study designs and implements a light-weight cache algorithm for ICN routers to achieve both light-weight computation and high cache hit rates. As the result, the study shows a software router architecture for high speed forwarding in a future Internet architecture.

研究分野：コンピュータネットワーク

キーワード：情報指向ネットワーク インターネット ルータ パケット転送 キャッシュ

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

情報指向ネットワーク(ICN)は次世代インターネットアーキテクチャの有力な候補であり、ルータには高性能と低消費電力の両立が求められている。これに対して、専用の LSI を用いたハードウェアルータは低消費電力化が困難であり、低消費電力な PC ベースのソフトウェアルータを ICN ルータに適用することが期待されている。しかしながら、現状の最速の技術を用いても約 20 ギガビット/秒のパケット転送速度であり、商用ルータに必要な最低限の約 320G ビット/秒(1000 万 PPS に相当)の達成には約 16 台の PC が必要で、ルータの容積が大きくなり過ぎ、普及の足かせとなっている。従って、ソフトウェア ICN ルータを 2 倍以上高性能化して PC 台数を半減することで、商用ハードウェアルータと同程度の容積にすることが重要な研究課題となっている。

2. 研究の目的

ICN は、キャッシュ、移動、マルチキャストを自然にサポートする、有力な次世代インターネットアーキテクチャであるが、これら豊富な機能を提供するルータの高性能化が普及の鍵である。一方、次世代インターネットには、増加一方の消費電力の削減が求められている。これに対して、本研究提案では、PC ベースのハードウェアプラットフォームを用いて、100 万 PPS(Packet Per Second)以上の高いパケット転送速度と、低消費電力を両立するソフトウェア ICN ルータアーキテクチャを、マルチコア CPU の並列性を最大限に活用することで実現する。これにより、インターネットへの ICN の導入を加速するとともに、社会的な要請となっているインターネットの消費電力削減に貢献する。

3. 研究の方法

マルチコア CPU をプラットフォームとするソフトウェア ICN ルータにおいて、高速で高いキャッシュヒット率を提供するパケット転送を実現するために、以下の 3 つの手法を開発する。理論的な検討だけでなく、プロトタイプを実装することで、高速なパケット転送と高いヒット率を実現することを実証する。

(1) 高速パケット転送を実現するプリフェッチ技術

ソフトウェアルータにおいて、高速なパケット転送を阻害する要因は、低速な DRAM にパケット転送用のデータ構造が記録され、パケット受信毎に DRAM から読み出すための待ち時間である。ICN では、宛先数が多くデータ構造の大半が DRAM に蓄積されることになり、待ち時間によるパケット転送が制限されることが課題である。これに対して、ICN のパケット処理の特性を活用して、データ構造を DRAM からプリフェッチすることで、待ち時間を無くし、理想的なパケット転送速度を実現する。

(2) 軽量のキャッシュ方式

ビデオなどの大容量のコンテンツが増加しており、ICN のキャッシュを活用したトラフィック削減が期待されている。一方、LFU などのキャッシュヒット率が高いキャッシュ方式、具体的には、キャッシュエビクションはパケット単位でキャッシュ処理を行うことは困難である。これに対して、頻度ベースのキャッシュアドミッションに基づく軽量のキャッシュ方式を開発する。

(3) CPU コアの負荷分散技術

(1)、(2)の手法を適用するマルチコア CPU において、受信したパケットが全ての CPU コアに分散されることを仮定して、プロトタイプによる評価を行ってきた。一方、現実には、均等な負荷分散は容易でなく、均等に負荷を分散すると、ステータフルなパケット転送を行う ICN では、ルータで排他制御を実施する必要が生じる。これに対して、不要な排他制御回数を削減し、CPU コアに受信パケットを均等に分散する手法を開発する。

4. 研究成果

研究方法に従って、マルチコア CPU を用いた PC プラットフォームとしたソフトウェアルータにおいて、高速なパケット転送と高いヒット率を提供するパケット転送方式を開発し、プロトタイプ実装することで性能を実証した。低消費電力な PC プラットフォーム上で実装することで、高速性と省電力を両立している。

(1) 高速パケット転送を実現するプリフェッチ技術

第一に、ICN ソフトウェアを PC ハードウェアプラットフォーム上で実行し、CPU 命令レベルの動作を解析し、キャッシュの制御用データ構造が蓄積された DRAM からのアクセス時間が性能向上のボトルネックになっていることを明らかにした。第二に、ボトルネックとなるアクセス時間を隠蔽して、パケット転送速度を向上させるように、複数のパケットを同時処理することで、制御用のデータ構造をプリフェッチする手法を設計した。設計した手法を、ICN ソフトウェアにプロトタイプ実装することで、68.3 G ビット/秒の転送速度を実現できることを実証した。既存手法と比較して、約 1.31 倍の転送速度

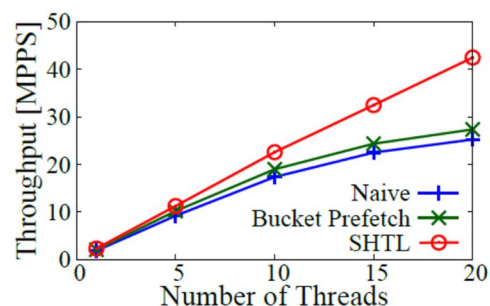


図1 パケット転送速度

を実現している。

提案したプリフェッチ手法は、ハッシュベースの packets 転送方式(図中の Naive)に、ハッシュの packets をプリフェッチする手法(図中の Bucket Prefetch), さらに Longest Prefix Match に対して、長さの短いプレフィックスの DRAM 上の FIB(Forwarding Information Base)から投機的にプリフェッチする手法(図中の SHTL)を加えて実現している。図 1 には、スレッド数、正確には、CPU コア数を増加させた時の各 packets 転送方式のスループット、具体的には packets 転送測度(メガ packets /秒)を示している。図 1 より、提案する packets 転送方式(SHTL)が、スレッド数の増加に比例して増加している。この研究成果は、英文論文誌 Computer Networks に、掲載された[]。

(2) 軽量のキャッシュ方式

高いヒット率と少ない計算量の頻度ベースのキャッシュアドミッション方式(フィルタと呼ぶ)を設計した。図 2 に示すように、過去 h 個分の受信 packets の名前 (x_i) を記録し、さらに名前毎の出現回数 (f_{x_i})、すなわち要求回数を記録する。新たな Data packets が到着した時、その名前の要求回数 (f_{x_i}) がしきい値以上の場合のみ、キャッシュ (CS と呼ばれる) に挿入する。また、削除する Data packets については、FIFO のような軽量のキャッシュエビクション方式を採用する。フィルタと FIFO エビクションを組み合わせたキャッシュ方式をマルコフ過程を用いてモデル化し、高いヒット率を実現することを検証した。この研究成果は、英文論文誌 IPSJ Journal に、掲載された[]。

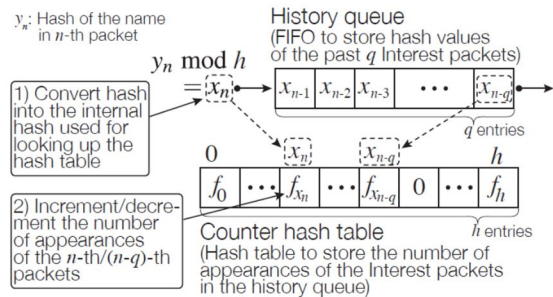


図 2 フィルタのデータ構造

また、ビデオのような大容量のコンテンツ(オブジェクト)が転送される場合、コンテンツを多数の packets (Data packets) に分割すると、先頭の packets がキャッシュから追い出されると、残りの packets がヒットしなくなり、キャッシュヒット率が低下する。Looped replacement と呼ばれるこの問題の発生原因を、図 3 に示すように、ICN の packets レベルのキャッシュの動作をマルコフ過程でモデル化することで明らかにした。具体的には、人気の低い Data packets が looped replacement が発生するコンテンツに割り込みことが発生原因であることを明らかにした。さらに、頻度ベースのキャッシュアドミッション方式が looped replacement を解析することを理論的に証明した。フィルタに関する研究成果は、英文論文誌 IPSJ Journal に掲載され[], looped replacement に関する研究成果は国際会議 ACM ICN2018 で発表した。

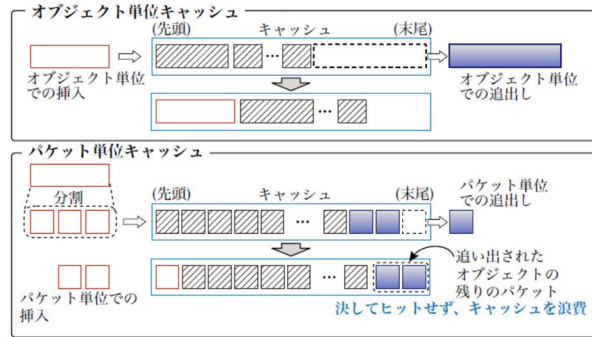


図 3 looped replacement の解析モデル

(3) CPU コアの負荷分散技術

ICN ルータは、ステートフルな packets 転送を行うために、フローの転送情報(フロー情報と呼ぶ)を Pending Interest Table (PIT) と呼ぶテーブルに転送情報保存する。このフロー情報は packets 転送ごとに書き換えられるため、ICN ルータを用いた PC プラットフォームで実行する場合、PIT 処理に排他制御が必要である。このため、PIT 内のフロー情報にアクセスする際は、Compare-and-Swap (CAS) 命令を用いて排他制御を行う必要がある。まず、第一に、同じ名前宛の packets を異なる CPU コアに割り当てて、常に、排他制御を行う条件で、22CPU コアのプラットフォームで packets 転送速度を測定した。図 4 に指名すように、CAS 命令を発行して排他制御を行う(w/locking)と、排他制御しない場合(w/o locking)しない場合と比較して、CPU コア数(スレッド数)が増加するにつれて、packets 転送速度が低下することを明らかにした。

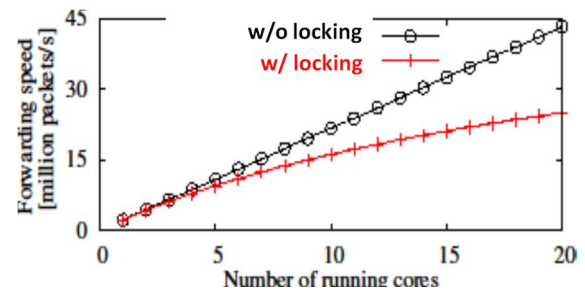


図 4 排他制御による packets 転送速度の劣化

第二に、この問題を解決するため、再送などのプロトコル処理によって回復可能なエラーは許容しつつ、排他制御を行わなくても、回復できないエラーが発生しない PIT を設計した。まず、PIT 処理で排他制御をしない場合、メモリリーク、ヌルアクセス、二重解放のプログラムエラーが発生し、プログラムの停止に直結する。いずれのエラーも動的確保された領域の操作によるものであるため、これらのエラーが発生しないように、領域の動的確保を用いないようにする。次に、排他制御をしない場合、異なる名前の複数の Interest packets に関する情報が同一の PIT エントリに書き込まれる可能性がある。一方が他方の情報を上書きする場合、上書きされた方の Data あるいは Nack は転送されない。複数の情報が PITに残った場合は、先に戻っ

てきた Data あるいは Nack が両方の Interest 受信フェイスに向けて転送される。つまり、想定していない Data あるいは Nack パケットが要求者に返送される。Data パケットについては、いずれ問題にもパケットの再送あるいは破棄

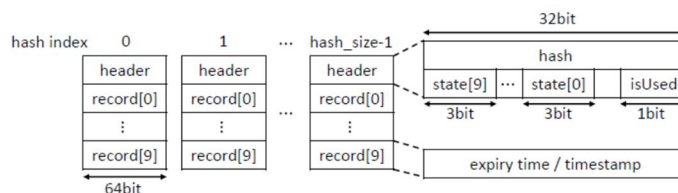


図 5 PIT のデータ構造

で対応できる。しかし後者のエラーが発生し誤った Nack が転送される場合は、誤った経路修正が発生し、これは ICN プロトコルで回復することはできないため、後者が発生しないように設計する。

具体的には、図 5 に示す PIT のデータ構造を設計した。まず、PIT はパケット転送のために書き換えが発生するため、挿入と削除処理が定数時間であるハッシュテーブルを採用する。エントリは、要素として名前のハッシュ値、受信情報、送信情報を保持する。受信情報として、Interest を受信したフェイス ID 等を保持する。第一の方針より、可変長の名前の代わりに、固定長の名前のハッシュ値を用いる。また、受信情報と送信情報は、Interest を送受信ごとに領域を確保するのではなく、ルータが有するフェイス数の要素をもつ静的な配列 record に格納し、フェイス ID をインデクスとしてアクセスする。第二の方針を満たすため、PIT エントリに関するメタ情報が必ず 1 つのスレッドによって書かれた状態を保つ。メタ情報として、1 つの変数 header に、名前のハッシュ値、record の各要素の状態、PIT エントリが使用されているかどうかの状態を 64 ビット変数にまとめる。これは、1 サイクルでストア命令の実行が可能であり、必ず最後に header を書き込んだスレッドの情報のみが採用される。

設計した PIT を採用したルータを想定して、2 つのパケットが同時に到着する場合に、PIT の転送情報が 2 つの CPU コアがアクセスする全ての順に対してパケット転送を行うシミュレータを作成し、ICN プロトコルで回復できないエラーが発生しないことを検証した。

(4) キャッシュ手法の調査

ユーザ間の関係性を表現したソーシャルネットワークを用いて、効率的なコンテンツキャッシュを実現する手法に関するサーベイを行なった。その結果、テンポラルネットワーク解析や多層ネットワーク解析の手法の利用が従来手法の改良に有望であることを明らかにした。サーベイ結果は、サーベイ論文として発表した[]。

< 引用文献 >

Junji Takemasa, Yuki Koizumi, and Toru Hasegawa, "Data Prefetch for Fast NDN Software Routers based on Hash Table-based Forwarding Tables," Computer Networks, Volume 173, 22 May 2020.

Junji Takemasa, Yuki Koizumi and Toru Hasegawa, "Lightweight Cache Admission Algorithm for Fast NDN Software Routers," IPSJ Journal of Information Processing, pp. 125-134, Feb. 2019

Yoji Yamamoto, Junji Takemasa, Yuki Koizumi and Toru Hasegawa, "Analysis on Caching Large Content for Information Centric Networking," in Proceedings of ACM Conference on Information-Centric Networking (ICN), Sept. 2018.

Sho Tsugawa, "A survey of social network analysis techniques and their applications to socially aware networking," IEICE Transactions on Communications, Vol.E102-B, No.1, January 2019.

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 3件）

1. 著者名 武政淳二、小泉佑揮、長谷川亨	4. 巻 27
2. 論文標題 Lightweight Cache Admission Algorithm for Fast NDN Software Routers	5. 発行年 2019年
3. 雑誌名 IPSJ Journal of Information Processing	6. 最初と最後の頁 125-134
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 津川翔	4. 巻 Vol. E102-B, No.1
2. 論文標題 A survey of social network analysis techniques and their applications to socially aware networking	5. 発行年 2019年
3. 雑誌名 IEICE Transactions on Communications	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transcom.2017EBI0003	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Junji Takemasa, Yuki Koizumi, and Toru Hasegawa	4. 巻 Vol 173
2. 論文標題 Data Prefetch for Fast NDN Software Routers based on Hash Table-based Forwarding Tables	5. 発行年 2020年
3. 雑誌名 Computer Networks	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.comnet.2020.107188.	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計10件（うち招待講演 0件/うち国際学会 2件）

1. 発表者名 山本瑠司、武政淳二、小泉佑揮、長谷川亨
2. 発表標題 Analysis on Caching Large Content for Information Centric Networking
3. 学会等名 ACM Conference on Information Centric Networking (ICN) (国際学会)
4. 発表年 2018年

1. 発表者名 山本瑠司、武政淳二、小泉佑揮、長谷川亨
2. 発表標題 ICN/ルータにおけるパケット単位キャッシュ挿入判別による大容量オブジェクトの自己置換の抑制
3. 学会等名 電子情報通信学会技術研究報告 (IN2018-73)
4. 発表年 2019年

1. 発表者名 山本瑠司、武政淳二、小泉佑揮、長谷川亨
2. 発表標題 キャッシュ挿入判別を用いた自己置換の抑制に関する一考察
3. 学会等名 電子情報通信学会ソサイエティ大会講演論文集 (B-7-10)
4. 発表年 2019年

1. 発表者名 津川翔
2. 発表標題 Empirical analysis of the relation between community structure and cascading retweet diffusion
3. 学会等名 International AAAI Conference on Web and Social Media (国際学会)
4. 発表年 2019年

1. 発表者名 津田 奈子、津川翔
2. 発表標題 ソーシャルネットワークのトラス構造がネットワーク上の情報拡散に与える影響
3. 学会等名 電子情報通信学会コミュニケーションクオリティ研究会
4. 発表年 2018年

1. 発表者名 津川翔
2. 発表標題 ソーシャルネットワークのコミュニティ構造がソーシャルメディア上の投稿の拡散規模に与える影響の分析
3. 学会等名 人工知能学会全国大会
4. 発表年 2018年

1. 発表者名 辻健太, 武政淳二, 小泉佑揮, 長谷川亨,
2. 発表標題 生存時間の短いコンテンツのLRUキャッシュヒット率への影響
3. 学会等名 電子情報通信学会ソサイエティ大会, B-7-30
4. 発表年 2017年

1. 発表者名 武政 淳二, 小泉 佑揮, 長谷川 亨
2. 発表標題 PC ベース NDN ルータのための高速なキャッシュアルゴリズムに関する一考察
3. 学会等名 電子情報通信学会総合大会, B-7-36
4. 発表年 2018年

1. 発表者名 山本瑠司, 武政淳二, 小泉佑揮, 長谷川亨
2. 発表標題 情報指向ネットワークにおける大容量コンテンツのパケット単位キャッシュのモデル化と解析
3. 学会等名 電子情報通信学会技術研究報告 (IN2017-122)
4. 発表年 2018年

1. 発表者名 武政淳二, 小泉佑揮, 長谷川亨
2. 発表標題 PC ベース NDN ルータのための高速なキャッシュアルゴリズムの解析
3. 学会等名 電子情報通信学会技術研究報告 (IN2017)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	津川 翔 (Tsugawa Sho) (40632732)	筑波大学・システム情報系・助教 (12102)	
研究分担者	小泉 佑揮 (Koizumi Yuki) (50552072)	大阪大学・情報科学研究科・准教授 (14401)	