様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月　 ７ 日現在

機関番号：１４４０１
研究種目：基盤研究(B)（一般）
研究期間：2020～2022
課題番号：２０Ｈ０４１７６

研究課題名（和文）クラスタ型NDNルータにおけるテラビット/秒高速パケット転送方式

研究課題名（英文）Tbit/s High Speed Forwarding for a Clustered NDN Router

研究代表者
長谷川　亨（Hasegawa, Toru）

大阪大学・大学院情報科学研究科・教授

研究者番号：７０５７６２６４

交付決定額（研究期間全体）：（直接経費）　　17,400,000円

研究成果の概要（和文）：次世代インターネットアーキクチャNamed Data Networking (NDN)に対して，パケット処理をプログラム可能なP4スイッチを用いてルータを設計し，10テラビット/秒のパケット転送速度を実現した．第一に，P4スイッチと接続したサーバが協調することで，100万以上の名前プレフィクスを高速に最長一意検索する手法を設計した．第二に，不要なデータパケットの転送をP4スイッチとサーバ間で削減する手法を設計した．これらの手法をプロタイプすることで，設計通りのパケット転送速度を実証した．

研究成果の学術的意義や社会的意義
６G時代に向けて，インターネットを流れるトラフィック量は今後も増加する一方である．一方，大容量の専用メモリとLSIを用いて大量トラフィックを処理する，現在の商用ルータでは，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．これに対して，本研究では，消費電力の少ないプログラマブルスイッチを用いてテラビット/秒のルータを開発した．開発のベースなるパケット転送技術は，省電力と高速性を両立するる次世代ルータに採用され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 Named Data Networking (NDN) router on a programmable switch, i.e., a P4 switch. The key techniques for high speed forwarding are two-folded: First, longest prefix matching of 10 million name prefixes is achieved by cooperation of a P4 switch and a server connected to it. Second, traffic reduction methods between the P4 switch and the server improves the forwarding speed in the case that data requests hit the cache. The study implements the prototype of NDN router based on the above two methods and validates that the prototype achieves more than Tera bit/s forwarding speed.

研究分野：情報ネットワーキング

キーワード：情報指向ネットワーキング　インターネット　ルータ　パケット転送　プログラマブルスイッチ

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

Content Centric Networking (CCN) / Named Data Networking (NDN)は，ルータに通信フローの状態を保持することで，キャッシュによるトラフィック削減やフロー単位の輻輳制御など，IP が提供できない機能を有する次世代インターネットアーキテクチャである．パケット転送をソフトウェアで処理する CCN/NDN ソフトウェアルータで，100 ギガ(G)ビット/秒のパケット転送速度を実現しつつあったが，バックボーンルータで使用するには，テラ(T)ビット/秒のパケット転送速度が要求されていた．一方，当時，パケット処理に特化した ASIC を搭載したプログラマブルスイッチ(以降，単に P4 スイッチと呼ぶ)が登場し，10T ビット/秒のパケット転送が必要なネットワーク装置に採用されつつあった．しかしながら，負荷分散装置などの通信フローの状態を持たないステートレスな通信を対象としており，CCN/NDN のように通信フローの状態を有するステートフル通信を P4 スイッチ上に実装することが，重要な研究課題となっている．

２．研究の目的

本研究の目的は，高負荷時にもパケット廃棄を発生させることなく，10T ビット/秒のパケット転送速度を提供する，CCN/NDN ルータを P4 スイッチと計算機を組み合わせて実現することである．10G ビット/秒クラスの CCN/NDN ルータでのフロー状態は百万程度であることが分かっており，その実現には，1 千万のフロー状態を 1 秒間に 2 億回更新することが必要である．この厳しい条件のもとで，大量のフロー状態の一貫性を保ちながら，名前ベースの最長一致検索や CCN/NDN 固有のキャッシュ処理を 10T ビット/秒以上のパケット転送速度を実現することを目指す．この結果，通信インタフェース毎に具備していた高消費電力なメモリ装置である Ternary Content Addressable Memory (TCAM)を，持たない P4 スイッチを採用することにより，ルータ，ひいてはインターネットの消費電力削減に貢献することが期待できる．

３．研究の方法

P4 スイッチと計算機を組み合わせて，10T ビット/秒以上のパケット転送速度を実現するために，（１）～（３）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を実証する．また，実証のために，10T ビット/秒での実証を実現するためのテスタを開発する．

（１）P4 スイッチと計算機を組み合わせたルータアーキテクチャ

P4 スイッチの ASIC に搭載された SRAM、TCAM の高速メモリの容量は少ないため，CCN/NDN の大量の名前プレフィクスを記録する Forwarding Information Base (FIB)を P4 スイッチに持たせることは困難である．これに対して，P4 スイッチの複数の計算機を組み合わせ，計算機の DRAM 上に NDN FIB を収容するアーキテクチャを設計する．最適なパケット転送速度を実現する P4 スイッチと計算機の機能分担を設計するとともに，設計したルータアーキテクチャに基づくプロトタイプを実装する．

（２）計算機への負荷分散技術

（１）で開発するアーキテクチャでは，　CCN/NDN における要求(interest)パケットを複数の計算機に送信して，最長一致検索などの CCN/NDN のプロトコル処理を計算機で実行する．10T ビット/秒の実現には，複数の計算機が必要であるため，P4 スイッチからこれらの計算機に均等に要求パケットを割り当てない場合，負荷の高い計算機で要求パケットの損失が発生する．これに足して，全ての計算機に均等に要求パケットを割り当てる負荷分散技術を開発する．

（３）高速キャッシュ処理技術

（１），（２）に基づく CCN/NDN ルータはキャッシュ機能を提供していないため，（１）のアーキテクチャを改良し，キャッシュ処理を行いながら，10T ビット/秒のパケット転送速度を実現する．計算機にキャッシュを置くため，P4 スイッチと計算機間に転送される応答(データパケット)の転送量を削減することが鍵となる．

（４）高速テスタ技術

（１）～（３）の技術を用いて，10T ビット/秒のパケット転送速度の見通しを得たが，従来の CCN/NDN ルータの試験に利用されてきた汎用計算機ベースのトラフィックジェネレータは高々100 Gbps の試験トラフィックしか生成できない。これに対し、P4 スイッチを用いて，T ビット/秒の試験トラフィックを生成する CCN/NDN ルータの試験用のテスタを開発する．

４．研究成果

研究方法に従って，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．

（１）P4 スイッチと計算機を組み合わせたルータアーキテクチャ

P4 スイッチ ASIC 単体では，SRAM, TCAM へ CCN/NDN の FIB を全て収容することは難しく、計算機単体では CPU の低速な処理により 100 G ビット/秒程度のパケット転送速度しか達成できない．このため，図 1 のように複数の計算機を P4 スイッチへ接続しルータハードウェアを構築する．ただし，計算機の接続にはスイッチのポートを消費し，外部とのパケット転送に必要なポート数が減少する．つまり，転送可能なパケット数を増加するために計算機の数を増加すると外部

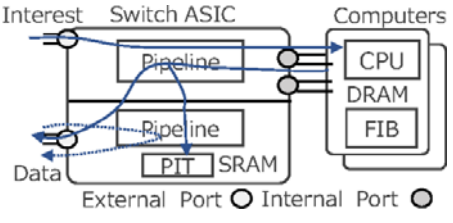

図 1 ルータのアーキテクチャ

ポート数，つまり外部と通信する帯域が減少する．したがって，計算機の接続ポート数を最小にし、外部ポートの帯域を確保するよう、パケット処理を設計する．CCN/NDN の通信は Interest と Data の要求-応答型であり、Data のサイズは Interest より十分大きいため，Data の処理を P4 スイッチだけで行うことで，必要な計算機数を削減する．具体的には PIT を P4 スイッチの SRAM，FIB を計算機の DRAM へ収容する．これにより FIB にアクセスする Interest は計算機で処理するが，PIT にしかアクセスしない Data はスイッチだけで処理することが可能になる．Data のサイズが Interest より大きいため、P4 スイッチと計算機間で消費する帯域を削減できる．

図 1 にパケット処理フローを示す．まず，P4 スイッチは Interest を受信すると計算機へ渡し，計算機は Interest の転送ポートを FIB で検索後スイッチへ渡す．次に，P4 スイッチは PIT へ Interest の受信元ポートを記録し，外部へ転送する．ここで，ASIC はパケットを処理するパイプラインを複数搭載しており，異なるパイプライン間の SRAM の読み書きを禁止している．このため，Interest の受信元ポートは戻りの Data が到着，つまり Interest を転送するパイプラインに記録する．一方，スイッチは Data を受信すると，PIT から Interest の受信元ポートを読み込んだ後に削除し，外部へ転送する。

表 1 パケット転送速度

Router Architecture	bits/s	packets/s
Proposal	470 Gbps	94.4 MPPS
Naive	79 Gbps	15.8 MPPS

プロトタイプを開発し，Tofino （商用の P4 スイッチ）と 22 コアの Xeon CPU を有する計算機を用いて，提案したルータ（Proposal）と全てのプロトコル処理を計算機で行うルータ（Naïve）の性能評価を実施した結果を表 1 に示す．表 1 では，P4 スイッチの 1 つのパイプラインを用いた場合のパケット転送速度を測定した．最近の Tofino-2 では 16 個のパイプラインを有しており，Tofino-2 を用いれば，数 T ビット/秒のパケット転送が可能になる．

この研究成果は，国際会議 ACM ICN 2021 で発表した[①]．

（２）計算機への負荷分散技術

既存の CCN/NDN ルータでは，複数の CPU コア，あるいは計算機で受信したパケットを割り当てる際に，同じ名前のパケットを同じ CPU コア/計算機に割り当てる負荷分散方式（シャーディングと呼ぶ，Sharding）を採用している．これによる排他制御（mutex）を排している。しかし，コンテンツの名前毎にパケットの到着数が偏る場合，CPU コア/計算機への割り当てパケット数が不均一になる．具体的には，インターネットのトラフィックで観測されるようにコンテンツの人気度に偏りがあると，高人気なコンテンツ名を扱う CPU コア/計算機へ計算能力を超えるパケット数が割り当てられ，パケットロスが頻発する．均一な割り当てを仮定した場合と比べ，全 CPU コア/計算機の計算能力を使い切れない分，パケット転送速度が低下する．

これに対して，高人気な名前のパケットを全 CPU コア/計算機に分散し，その他をシャーディングで割り当てることで，均一な割り当てを実現しつつ PIT の排他制御を回避する負荷分散技術を開発した。理想的な負荷分散（ideal）、提案した負荷分散術（proposal）、従来のシャーディング（sharding）と排他制御（mutex）のパケット転送速度とパケット損失率をシミュレーションで測定した．図 2 に示す通り，提案した負荷分散技術は，従来のシャーディングと排他制御と比較して高い性能を示している．

この研究成果は，論文誌 IEEE ACCESS （②） ならびに電子情報通信学会和文論文誌 （③） に採択されている．

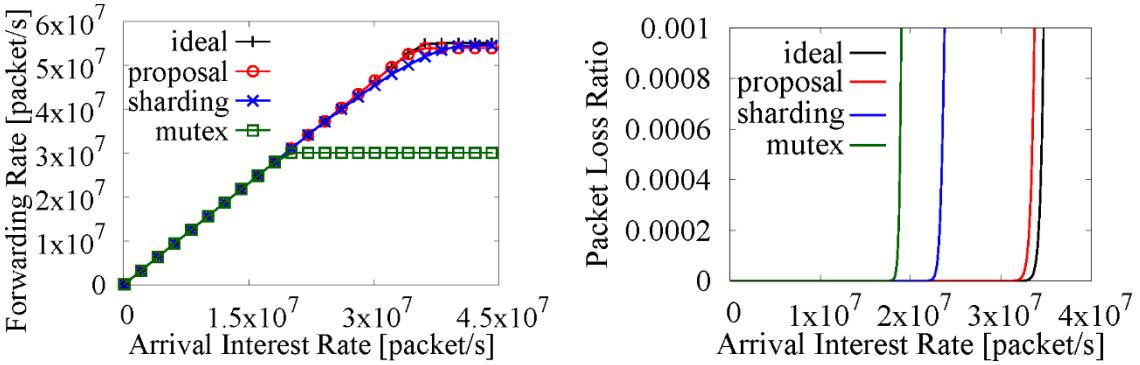

図 2 負荷分散技術の比較

（３）高速キャッシュ処理技術

P4 スイッチと計算機で構成する CCN/NDN ルータでキュッシュ処理を行う場合の，パケット転送速度を律速する要因を明らかにするとともに，律速要因の解消法を設計した。律速要因の解析では，P4 スイッチの ASIC，計算機の CPU，双方を繋ぐポートの負荷より導出したパケット転送速度を比較し，P4 スイッチと計算機を繋ぐポートの帯域が律速要因であることを明らかにした．

この結果より，キャッシュに起因するデータを伝送するために計算機をつなぐポートの数を削減し，かわりに外部とのパケット転送に必要なポート数を増加することが高速化に必要であり，このために，P4 スイッチと計算機間のデータ伝送の削減が必要となる示唆を得た．

これに対して，律速要因の解消法として，P4 スイッチから計算機方向へのデータ伝送，逆方向のデータ伝送を削減する方式をそれぞれ設計した．まず，P4 スイッチから計算機へのデータ伝送は，新たに到着するコンテンツをキャッシュへ挿入することに起因するため，将来ヒットしないコンテンツをスイッチでフィルタするキャッシュアドミッションを設計した．次に，計算機から P4 スイッチへのデータ伝送は，ヒットした要求に対する応答のコンテンツをキャッシュから読み出す処理であり，同じコンテンツへの複数要求を一定時間遅延させ，1 度の読み込みにより一括で応答する遅延応答を設計した．遅延応答の詳細を，図 3 に示す．

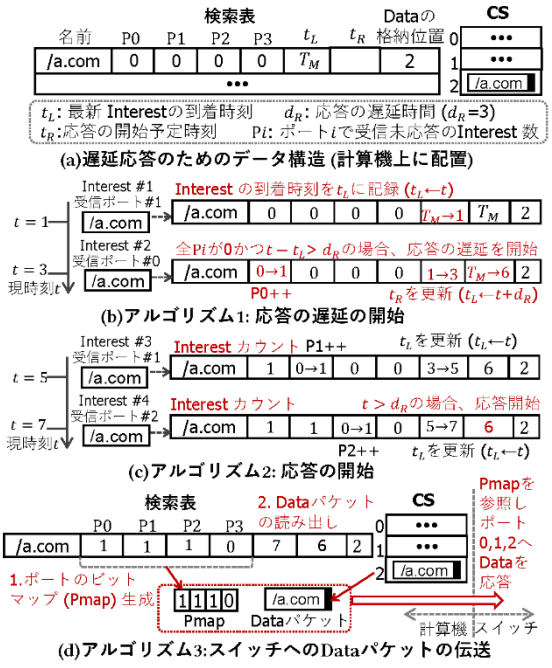

図 3 遅延応答

提案方式を評価するため， Tofino スイッチと 2 台の汎用計算機上に実装し，提案した手法を実装したルータ(提案ルータ)と実装しないルータ(従来ルータ)の，キャッシュ処理時のパケット転送速度を測定した．この結果，表 2 に示すように，従来の 543G ビット/秒から 916 G ビット/秒にパケット転送速度を向上させた．表中の $b_{s\to c}$ と $b_{c\to s}$ は，P4 スイッチから計算機，および計算機から P4 スイッチのデータパケットのトラフィック量を示しており，提案手法が削減していることが分かる．

表 2 パケット転送速度

	m_c [ポート]	$b_{s\to c}$ [Gbps]	$b_{c\to s}$ [Gbps]	フォワーディング速度 [Gbps]
提案ルータ	4	242	301	916
従来ルータ	4	399	285	543

この成果は，国際会議 Global Internet Symposium (④) に発表した．

（4）高速テスタ技術

P4 スイッチは，100 G ビット/秒 のポートを数 10 ポート，ならびに，これらのポートに対しトラフィックを生成可能なハードウェア生成器 pktgen を備えており，10T ビット/秒級の試験を単一のスイッチハードウェアを用いて安価に構築できる．しかし，CCN/NDN のトラフィックを生成するには，以下の 2 つの課題がある．第一の課題は，要求-応答の双方向通信を模擬することである．しかし，pktgen 自身は一定の時間間隔でパケットを生成する機能しか有していない．第二の課題は、多種類のコンテンツ名を要求する Interest の系列を生成することである。CCN ルータの適用先として有望であるインターネットでは、10 億種類程度の異なるコンテンツ名を要求する Interest のトラフィックが想定されている．これに対し、pktgen では生成するパケットのパターンとして，高々 250 種類程度の異なる名前のパケットしか定義できない．

これに対して，2 つの課題を解決し，約 10 億のコンテンツ名から成る要求-応答パケットの組を 10T ビット/秒の速度で生成可能な CCN トラフィックジェネレータ ccnGen を開発した．まず，第一の課題に対し，pktgen を用いた Interest の生成，CCN ルータから受信したコンテントオブジェクト の返送をスイッチ ASIC の 2 つのパイプラインを用いてそれぞれで行う要求-応答型の双方向トラフィック生成器を設計した。次に，第二の課題に対し，pktgen において名前を除く Interest のヘッダを有するパケット(テンプレート Interest) のパターンを 1 つだけ定義し，ASIC のパイプラインで異なる名前を付与することで，Interest に付与する名前の種類を増幅する．このために，約 10 億種類の異なる名前を小容量の SRAM メモリを用いて生成可能な ASIC パイプラインの構築法を設計した．図 4 に ccnGen の構成を示す．

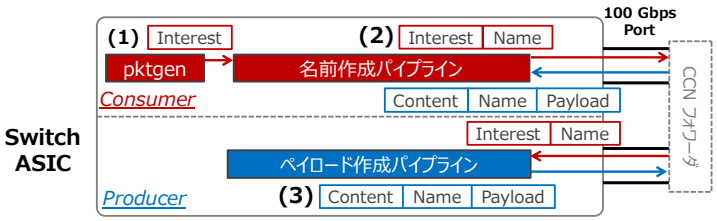

図 4 ccnGen の構成

1 台の Tofino スイッチで実験した結果，1.6T ビット/秒で試験トラフィックを生成できることを確認した．さらに，ccnGen は（1）、（3）のルータのプロトタイプの評価実験に使用した．

この成果は，国際会議 ACM ICN のデモ/ポスター論文（⑤）として，発表した．

さらに，通信プロトコルのトラスト分析を主な適用先とする，動作仕様のトレース包含やトレース一致についての検討を行い，効率的な検証手法を開発した．この結果は，ccnGen を用いた試験シナリオの作成に活用した．

＜引用文献＞

① Junji Takemasa, Yuki Koizumi, Toru Hasegawa, “Vision: Toward 10 Tbps NDN Forwarding with Billion Prefixes by Programmable Switches,” in ACM Conference on Information-Centric Networking (ACM ICN 2021), Sep. 2021.

② Junji Takemasa, Atsushi Tagami, Yuki Koizumi, Toru Hasegawa, “Load Balancing for Stateful Forwarding by Mitigating Heavy Hitters: A Case for Multi-Threaded NDN Software Routers,” IEEE Access, vol. 8, pp. 155071-155085, 2020.

③ 武政淳二, 小泉佑揮, 長谷川亨, “汎用計算機ベースの高速な情報指向ネットワークルータの実装,” 電子情報通信学会論文誌 B, Vol.J106-B, No.5, pp.265-280, May 2023.

④ Junji Takemasa, Yuki Koizumi, Toru Hasegawa, “Terabytes and Terabits/s Packet Caching in ICN Routers using Programmable Switches,” in IEEE Global Internet Symposium 2022 (in conjunction with IEEE CloudNet 2022), pp. 67-72, Nov. 2022.

⑤ Junji Takemasa, Ryoma Yamada, Yuki Koizumi, Toru Hasegawa, “ccnGen: A High-speed Generator of Bidirectional CCN Traffic Using A Programmable Switch,” in ACM Conference on Information-Centric Networking (ACM ICN 2021), Poster/Demo Session, Sep. 2021.

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　2件）

１．著者名 武政淳二、田上敦士、小泉佑揮、長谷川亨	４．巻 vol. 8
２．論文標題 Load Balancing for Stateful Forwarding by Mitigating Heavy Hitters: A Case for Multi-Threaded NDN Software Routers	５．発行年 2020年
３．雑誌名 IEEE Access	６．最初と最後の頁 155071, 155085
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2020.3018555	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 武政淳二、小泉佑揮、長谷川亨	４．巻 vol. 173
２．論文標題 Data prefetch for fast NDN software routers based on hash table-based forwarding tables," Computer Networks	５．発行年 2020年
３．雑誌名 Computer Networks	６．最初と最後の頁 107188
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 －

１．著者名 武政淳二，小泉佑揮，長谷川亨	４．巻 J106-B
２．論文標題 汎用計算機ベースの高速な情報指向ネットワークルータの実装	５．発行年 2023年
３．雑誌名 電子情報通信学会論文誌B	６．最初と最後の頁 265-280
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transcomj.2022NSI0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計13件（うち招待講演　1件／うち国際学会　2件）

１．発表者名 武政淳二、小泉佑揮、長谷川亨
２．発表標題 Vision: Toward 10 Tbps NDN Forwarding with Billion Prefixes by Programmable Switches
３．学会等名 Proceedings of ACM ICN 2021（国際学会）
４．発表年 2021年

１．発表者名
武政淳二、山田涼真、小泉佑揮、長谷川亨

２．発表標題
ccnGen: A High-speed Generator of Bidirectional CCN Traffic Using A Programmable Switch

３．学会等名
Proceedings of ACM ICN 2021（国際学会）

４．発表年
2021年

１．発表者名
長谷川亨

２．発表標題
Revisiting High-Speed Forwarding: Cases for State-full forwarding and Learned Index based Forwarding

３．学会等名
1st International Workshop on Theory and Practice of Programmable Forwarding（招待講演）

４．発表年
2021年

１．発表者名
武政 淳二，田上 敦士，小泉 佑揮，長谷川 亨

２．発表標題
プログラマブルスイッチと計算機を組み合わせた10TbpsのNDNパケットフォワーディングに関する一考察

３．学会等名
電子情報通信学会技術研究報告

４．発表年
2021年

１．発表者名
山田 諒真，武政 淳二，小泉 佑揮，長谷川 亨

２．発表標題
P4プログラマブルスイッチを活用したCCNトラフィックジェネレータ

３．学会等名
電子情報通信学会 第19回ICN研究会ワークショップ

４．発表年
2021年

１．発表者名
武政 淳二，小泉 佑揮，長谷川 亨

２．発表標題
プログラマブルスイッチと汎用計算機を組み合わせた NDN ルータに関する一考察

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

４．発表年
2021年

１．発表者名
山田 諒真，武政 淳二，小泉 佑揮，長谷川 亨

２．発表標題
プログラマブルスイッチを用いた CCN の双方向トラフィック生成に関する一考察

３．学会等名
電子情報通信学会技術研究報告

４．発表年
2021年

１．発表者名
河辺義信

２．発表標題
二次元的トラスト表現法の大学生の就職支援への適用

３．学会等名
日本知能情報ファジィ学会　ソフトサイエンス研究部会　第３２回ソフトサイエンス・ワークショップ

４．発表年
2022年

１．発表者名
小山 亮，武政 淳二，小泉 佑揮，田上 敦士，長谷川 亨

２．発表標題
排他制御しない PIT を用いたマルチコア NDN ルータで発生するパケット転送誤りからの回復に関する一考察

３．学会等名
電子情報通信学会技術研究報告

４．発表年
2020年

１．発表者名
小山 亮，小泉 佑揮，長谷川 亨

２．発表標題
排他制御しない PIT を用いたマルチコア NDN ルータで発生するエラーの検証法に関する一考察

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

４．発表年
2020年

１．発表者名
武政 淳二，田上 敦士，小泉 佑揮，長谷川 亨

２．発表標題
高人気パケットの分散割り当てによるマルチスレッドNDNルータの高速化

３．学会等名
電子情報通信学会第 18 回情報指向ネットワーク研究会

４．発表年
2020年

１．発表者名
武政 淳二，田上 敦士，小泉 佑揮，長谷川 亨

２．発表標題
高人気パケットの分散割り当てによるマルチスレッドNDNソフトウェアルータの高速化に関する一考察

３．学会等名
電子情報通信学会技術研究報告

４．発表年
2020年

１．発表者名
河辺義信

２．発表標題
トラスト遷移の検証のためのシミュレーション関係の自動生成

３．学会等名
日本知能情報ファジィ学会第31回ソフトサイエンス・ワークショップ

４．発表年
2021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	小泉　佑揮 (Koizumi Yuki) (50552072)	大阪大学・大学院情報科学研究科・准教授 (14401)	
研究分担者	河辺　義信 (Kawabe Yoshinobu) (80396184)	愛知工業大学・情報科学部・教授 (33903)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関