

光 L2 網におけるマルチドメインに対応するアドレス割り当ての検討

鍵田 幸大[†] 佐藤 丈博[‡] 岡本 聡[†] 山中 直明[†]

[†] 慶應義塾大学 大学院理工学研究科 〒223-8522 神奈川県横浜市港北区日吉 3-14-1

[‡] 京都大学 大学院情報学研究科通信情報システム専攻 〒606-8501 京都府京都市左京区吉田本町 36-1

E-mail: [†] kodai.yarita@yamanaka.ics.keio.ac.jp

あらまし 近年の通信トラフィック大容量化及び、通信サービスの多様化に対応するため、フォトニックネットワーク及びノードの進展が求められており、解決策としてスマートフォトニックネットワークワーキング(SPN)が提案されている。SPN では、トランスポートに特化し、広域、マルチドメインに対応した新たなレイヤ 2(光 L2)が求められている。光 L2 においては、マルチドメインの一体運用、アドレス集約化によるルーティングの簡易化が求められており、新たなアドレス体系の構築やブロードキャストを使用しないアドレス解決を行うことが求められる。本稿では光 L2 網におけるマルチドメインに対応するためのアドレス割り当てについて検討した結果を報告する。

キーワード スマートフォトニックネットワークワーキング, 光 L2, アドレス割り当て, プログラマブル光エッジ

Design of Address Assignment for Multi-Domain Photonic L2 Network

Kodai YARITA[†] Takehiro SATO[‡] Satoru OKAMOTO[†] and Naoaki YAMANAKA[†]

[†] Graduate School of Science and Technology, Keio University 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, Kanagawa, 223-8522 Japan

[‡] Graduate School of Informatics, Kyoto University 36-1 Yoshida-honmachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan

E-mail: [†] kodai.yarita@yamanaka.ics.keio.ac.jp

Abstract In increasing communication traffic and diversifying services, progress of photonic networks and nodes are required. As a solution to that, Smart Photonic Networking (SPN) has been proposed. In the SPN, A new layer 2 specialized for transport and correspond to wide area and multi-domains is required, that is called Photonic Layer 2. For the Photonic Layer 2, it is required to simplify routing by integrating multi-domains and address aggregation. Therefore, it is necessary to construct a new address structure and resolve address without using broadcast. In this paper, we report the result of the address structure design for supporting multi-domains in the Photonic Layer 2

Keywords Smart Photonic Networking, Photonic Layer 2, Address Assignment, Programmable Photonic Edge

1. 背景

1.1. スマートフォトニックネットワークワーキング(SPN)

通信トラフィックの増大及び、通信サービスの多様化に対応するためのフォトニックネットワーク技術及びノード技術の進展が求められており、2020 年以降のフォトニックネットワークをターゲットとしたフォトニッククラウドが提案されている[1].

フォトニッククラウドでは、サービスごとにネットワークリソースを割り当てるスライスを構築するネットワーク仮想化[2]が重要な概念として採用されており、マルチプロトコル/マルチレイヤ Software Defined Networking (SDN)サービスネットワークを構築する Software Defined Transport Network (SDTN)[3]の実現が期待されている。SDN はデータコムを対象としたレイヤ 3 (L3)/レイヤ 2(L2) 制御技術であるが、SDTN では

通信キャリアの利用を対象としたマルチプロトコル/マルチレイヤ SDN であり、主にレイヤ 2 以下を対象としている。SDTN における要求はスライスの要求条件と相通じるものが多く、解決策としてスマートフォトニックネットワークワーキング(Smart Photonic Networking: SPN) が提案されている[4].

SPN では光技術を活用し SDTN に特化した、広域、マルチドメインに対応する新たなレイヤ 2 が求められており、そのプロトコルが光 L2 (Photonic L2:PL2) と名付けられている[1,4]. SPN の目標は、光技術に基づいたパケット+伝達網サービス基盤の実現であり、光 L2 を用いたマルチサービスの提供、サービスごとのスライスである光 L2 サービススライス提供を行うことである。SPN の実現にあたっては、フォトニックネットワーク処理用の各種モジュールを光インターポーザ

で結合して処理機能チップ・ボード・ノードを統一的に構築するフォトニックネットワークプロセッサ[1, 4]が提案されている。SPN の実現に向けた第一歩として、電子デバイスのスイッチで各種処理機能を結合して処理機能チップやボードをリソースプール化する、仮想通信処理プロセッサ (Virtual Reconfigurable Communication Processor: VRCP[5])が提案されている。光 L2 サービススライスをはじめとする SPN マルチサービスの提供には、STDN に適応する光ネットワークの終端装置であるプログラマブル光エッジ (Programmable Photonic Edge: PPE)[4, 5]を用いる。また光 L2 伝達網におけるスイッチングには、光 L2 フレームに基づきスイッチング (ブリッジ機能の提供) を行う光 L2-MAC-B[4, 5]ノードを用いる。光 L2-MAC-B に関しては、電気ラベルスイッチ・光ラベルスイッチの両方が想定されている [4, 6]。

伝達網においてレイヤ 1/レイヤ 2/レイヤ 3 が担う技術と既存技術がカバーしている領域、および光 L2 がカバーすべき領域を表 1 に示す。

表 1.各種トランスポート技術とカバー領域

技術名	網アドレス	ルーティング	リンク確立	エラー検出	データ送出	L4 収容	L3 収容	L2 収容	L1 収容
IP	◎	◎	○	×	×	◎	◎	◎	○
MPLS	×	○	◎	×	×	○	◎	◎	○
ATM	×	○	◎	◎	△	○	◎	◎	○
HDLC	×	×	◎	◎	×	○	◎	◎	○
Ethernet	○	△	◎	◎	◎	○	◎	◎	○
SDH/OTN (光L1)	×	×	○	△	◎	×	×	○	◎
光L2	△	○	◎	◎	◎	○	◎	◎	○

トランスポートプロトコルとしては Ethernet が多くの領域をカバーしているが、これを光 L2 に直接適用するためには、アドレス集約が困難、ブロードキャストを多用する [4, 5] という課題がある。前者に対しては、Ethernet アドレスの前半部を自由化する方式 [6, 7] が、後者に対しては分散ハッシュテーブルを用いた方式 [8, 9] が提案されている。

本論文の構成は以下の通りである。2 章では光 L2 フレームのフォーマットについて説明する。続く 3 章では光 L2 フレームのアドレスフィールドの割り当てを示し、4 章ではその光 L2 のアドレス割り当てについて説明する。最後に 5 章でまとめを述べる。

2. 光 L2 フレームフォーマット

電気ラベルを用いた光 L2 プロトコルでは、イーサネットに対する高可用性、IP との親和性、C-Plane の活用が求められる。そのため、マルチドメインの一体運

用、アドレス集約化によるルーティングの簡易化が求められており、新たなアドレス体系の構築やブロードキャストを使用しない技術が必要とされている [4, 5]。

また、ユーザ信号の光 L2 への収容(アダプテーション)手法においては、①アドレッシング自由度の確保、②ユーザ信号の透過性、③データグラム・パス及び、connection oriented なパスサービスの実現性、④C-Plane との親和性、の 4 要件が重要となる。これより、ユーザ信号全体を光 L2 フレームのペイロードデータとし、独立する光 L2 フレームヘッダを付与する独立マッピング方式の採用が適当であると考えられる。

以上の要件を満たすものとして、電気ラベル版光 L2 フレームフォーマットを提案した [6]。

光 L2 フレームを End-to-End で転送する上では、以下の 5 項目が重要となる。①宛先識別子：光 L2 フレームの宛先、②始端識別子：光 L2 フレームの送信元、③インスタンス 識別子：スライス・データフローの識別、④経路・キュー 識別子：ノードでの処理識別、⑤サービスタイプ識別子：光 L2 が提供するサービス識別

光 L2 フレームフォーマットは、これらを踏まえイーサネットとの整合性を重視し、必要最小限のフォーマット変更とした。図 1 に電気ラベル版光 L2 フレームフォーマット [6]を示す。

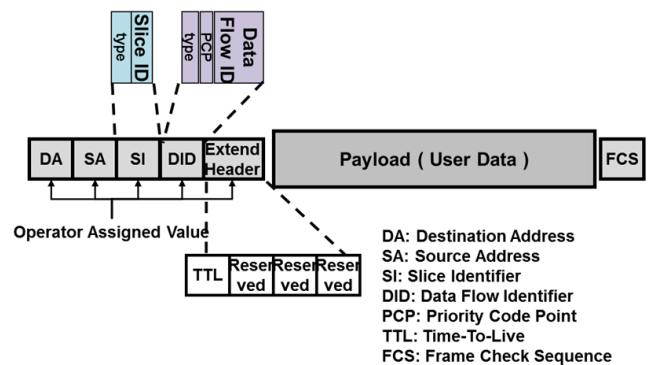


図 1.電気ラベル版フレームフォーマット

電気ラベル版光 L2 フレームフォーマットは以下の五つのフィールドから構成される。

- ① Destination Address (48 bit)
 - ② Source Address (48 bit)
 - ③ Slice Identifier (32 bit): Slice Type 16 bit, Slice ID 16 bit
 - ④ Data Flow Identifier (48 bit): Type 16 bit, Priority Code Point 8 bit, Data Flow ID 24 bit
 - ⑤ Extend Header (32 bit): TTL 8 bit, Reserved 24 bit
- 本稿では、① Destination Address (48 bit)及び、② Source Address (48 bit) フィールドにおいて、各 PPE の

通信処理性能, 保有するユーザ数, 光 L2 が支配するネットワークの大きさから, 状況に応じた柔軟なアドレス割り当て, アドレス集約化が可能となる, マルチドメインに対応するためのアドレス割り当てについて検討した結果を報告する.

3. 光 L2 アドレスフィールド

アドレスフィールドの 48 bit という値は, Ethernet における MAC アドレスに対応づけることができる. Ethernet の MAC アドレスにおいては, 上位 24 bit をベンダ ID, 下位 24 bit を Host ID として割り当てを行っているが, 光 L2 網においては, アドレス集約化の実現及び, IP ベースのコントロールプレーンとの融合を最大限に活用することを目指し, MAC アドレスに外部から自動配布されるノード ID とノード内で自動生成する物理的なポート番号(ノード内識別子)の組み合わせにより自動生成を行うものとする. よって, 上位 32 bit にノード ID (Domain ID, Region ID, Host ID)を与え, 下位 16 bit にノード内識別子を与える. 光 L2 網内では, 光 L2-MAC-B のブリッジノード[4, 5]を使用して, 上位 32 bit を利用したルーティングにより光 L2 フレームの転送が行われる.

光 L2 端点における宛先アドレス入手は下位 16 bit を指定しない Unnumbered 指定の他, 48 bit アドレスを指定する Numbered 指定の両方を想定する.

ノード ID について述べる. 先頭 2 bit はイーサネットと同様, ユニキャスト, マルチキャストの識別に用いる. イーサネットでは, 伝送時に 8 bit 単位で LSB→MSB と送信されるため, 内部表現(In the memory : I/G (Individual/Group) ビットは 8 bit 目, U/L (Universal/Local) ビットは 7 bit 目)と外部表現(On the wire : I/G は 1 bit 目, U/L は 2 bit 目)が異なっているが, 光 L2 では両者を外部表現で統一して扱う[6, 7]. また 2 bit 目については, 光 L2 では 1 としてイーサネットの Local アドレスの領域を利用することで, 既存 Ethernet での併用も可能とする. 以降の 30 bit を本稿では Domain ID, Region ID, Host ID としてアドレス割り当てを検討する.

図 2 に本稿で述べるアドレスフィールド(48 bit)の割り当て結果を示す.

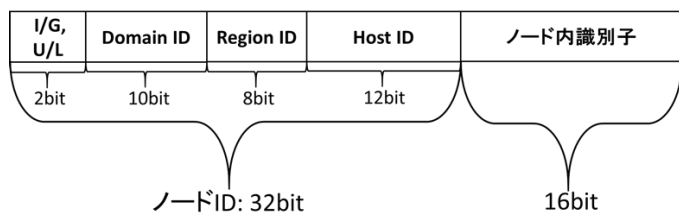


図 2. 光 L2 におけるアドレスフィールド

4. アドレス割り当て

4.1 Host ID フィールド

始めに Host ID について検討を行う.

PPE に導入を検討している VRCP の通信処理モジュール性能は 400 Gbps であり[5], 加えて PPE に複数の通信処理モジュールを搭載することで, ノードあたりの総入力インタフェース 3.2 Tbps, 総出力インタフェース 3.2 Tbps となる交換処理実現を目指している. ここで, 各 PPE が配下に保有するユーザ数について検討を行う. 前提として, FTTH (Fiber to the Home)を想定し, ユーザは複数の多重化段階を介し, 最終的に VRCP の入力インタフェースと通信を行うものとする.

現在の一般家庭における光回線は 1 Gbps (GE-PON) もしくは 2.5 Gbps (G-PON)の光回線を, 光スプリッタにより分割し, 複数のユーザで使用している. 光スプリッタでは最大 32 分割し供給しているとされているが, 32 分割上限まで供給した場合それ以上のユーザの追加ができないため, 実際には 16 分割程度での供給が妥当な割り当てといえる. また, VRCP の 400 Gbps と, PON の OLT (Optical Line Terminal)の出力 1 Gbps の間には, 10 Gbps, 100 Gbps のルータ等による多重化が存在すると想定した. 図 3 に, 1 台の PPE が収容するユーザ数の例を示した.

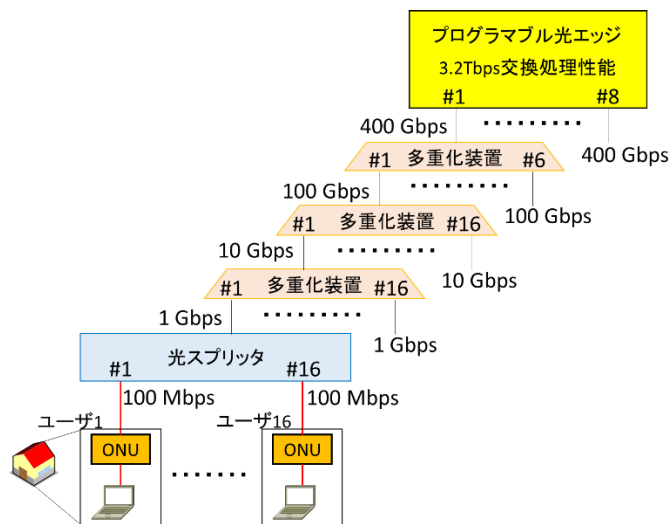


図 3. PPE の 1 インタフェースが収容可能なユーザ数

現在は 1 ユーザあたりの平均ダウンロードトラフィックは, 247.3 kbps であるが年率 37%で増加している[10]. GE-PON における保証レートは数 10 Mbps とされているが, 本稿では 1 ユーザあたり 100 Mbps を割り当てるとする. この場合, 全ユーザが同時に使用した場合の最大通信容量は 1.6 Gbps (16 ユーザを仮定)となる. しかし実際には全ユーザが同じタイミングで通信を要求することは想定しがたく, 半分の 0.8 Gbps を実際に

想定しうる最大通信容量とする。ここまでのユーザ(計 16 ユーザ)をユーザ群とし、同様のユーザ群が複数存在するとして、さらに上位に 10 Gbps のアップリンクを持つ多重化装置(IP ルータ等)を設けるとすると、それらのユーザ群×16 程度なら収容可能と想定できる。これに関しても 16 のユーザ群あたりの総最大通信容量は、 $0.8\text{Gbps} \times 16(\text{ユーザ群}) = 12.8\text{Gbps}$ となり、上記と同様に 10 Gbps の多重化装置で十分処理可能であると考えられる。さらに上位に、100 Gbps のアップリンクを持つ多重化装置を導入した場合、同様に 16 ユーザを保持するユーザ群(16 ユーザ)×16×16 まで収容可能と想定される。これらの上位に多重化装置が存在し、それが PPE の 400 Gbps の入力インタフェースと通信を行うものとする。この場合、上記と同様の理由で、最も上位の多重化装置では、最大通信容量として 100 Gbps が想定される光回線を、6 本収容することが妥当と言える。

以上の光回線の構成を見込むと、400 Gbps の通信処理モジュール性能を持つ入力インタフェース 1 つにつき、 $16 \times 16 \times 16 \times 6 = 24,576$ のユーザが収容可能といえる。また PPE では複数の通信処理モジュールを搭載することにより、3.2 Tbps までの通信処理性能実現を目指しているため、各 PPE が配下に持つ総ユーザ数は、 $24,576 \times 8 (3.2\text{Tbps}/400\text{Gbps}) = 196,608$ ユーザと考えられる。

ここで日本の総住宅戸数はおよそ 60,629,000 戸[11]であるため、すべての住宅の光回線を PPE で行うものとする、日本国内では $60,629,000 \div 196,608 \approx 308$ 台の PPE が必要となる。

ここでユーザ定義の範囲として、一般家庭以外にも企業が使用する ONU も追加する。日本における中小企業以上の会社はおよそ 382 万社[12]であり、従業員数は 4794 万人である。ここで従業員一人あたりに ONU を与えたとしても、国内で必要となる PPE の数は倍まで増えないうえ、現実的にそのような状況は考えにくい。Host ID 内でも余裕ができ、柔軟なアドレス割り当てを行い、アドレス集約化が可能となる。

また仮に光 L2 ネットワークの海外への拡張を考慮した場合、最も人口が多い中国を想定すると、総世帯数は約 4 億 3,000 万世帯[13]であるため、必要 PPE 数は $430,000,000 \div 196,608 \approx 2,187$ 台となる。

以上より、今後の拡張性も考慮し、アドレスフィールドの Host ID には 12 bit (4,096)を与えることが妥当であるといえる。12 bit を Host ID に与えることで、国内においては全住宅、中小企業の全従業員をユーザとして考慮した場合にも PPE に割り当てるアドレスフィールドは十分足りるといえ、海外をターゲットに見据えることも可能となる。

4.2 Region ID フィールド

続いて Region ID が使用するフィールドについて検討を行う。

Region ID には 8 bit を与える。一つの Region の単位を都市、都道府県レベルと考慮すると、日本では 47 都道府県、つまり 6 bit を Region ID として与えることで 64 以下の Region を登録することが可能となり、要件を満たす。ここで海外における光 L2 アドレスの汎用性を考慮すると、アメリカでは、州レベルの分割が可能となる。また、中国では都市レベルで分割を行った場合、人口 2,000 万人以上の大都市が 2 都市、人口 1,000 万人以上の中大都市が 13 都市、東京のおよそ半分にあたる人口 470 万人以上の中都市が 100 個存在している。ここで 7 bit を割り当てることにより、Region として 128 の分割が可能となり要件を満たすが、8 bit を Region ID として与えることで、主要都市内でさらに柔軟な割り当てを行うことを可能とする。また Region ID に 8 bit を与えた場合、日本においては必要数のおよそ 6 倍の猶予があり、必要に応じて県内でさらに 6 つ以上の Region を割り当てることのできる。これにより海外へのゲートウェイとなる PPE には専用の Region を割り当てる等が可能となり、役割に応じて都市、県内でより柔軟に割り当てを行うことを可能とする。

4.3 Domain ID フィールド

続いて、Domain ID が使用するフィールドについて検討する。光 L2 アドレスは、光 L2 網内で一意であれば良く、Global IP アドレスのように、全世界において一意である必要は無い。そのため Domain ID フィールドは、光 L2 網管理者が自由に利用することが可能である。しかしながら、Verizon のように地球規模の通信キャリアがルータ無で光 L2 網を利用した IP サービスを展開することも想定して、検討を行った。実際はアドレス解決に要する時間が距離遅延に支配され、中規模の光 L2 網をコネクションレスサービス向けに構築することは非現実的である。

Domain ID には 10 bit を割り当てる。最も大きなドメインとして、国を単位と想定する。ここで現在国際連合加盟国は 195 か国[14]であり、最低 8 bit、つまり 256 個分を Domain ID として与えることで要求は満たす。しかし、その場合国家内では Region ID 以降のアドレスフィールドのみから割り当てる必要が出てくる。ここで、Domain ID を 10 bit とすることで、各国平均およそ 5 つのドメインを使用することが可能となる。ただし各国 5 ドメインというのはあくまで全国家が光 L2 を導入した場合を仮定しており、現実的でない。つまり先進国をはじめとする主要国家にドメインを集中することが現実的であり、全世界で 1,024 のドメインを

扱うことが可能である。ここで契約者数 3,000 万人を超える世界の主要通信事業者数は 50 程度であり、国単位ではなく各国の通信事業者単位での割り当てにも対応可能となるため、光 L2 のターゲットとして世界への拡張も十分に可能であるといえる。

また国家間における通信には、光 L2-MAC-B を用いることにより、集約化された光 L2 アドレスによるルーティングが行われ、所望の光 L2 網内とつながる光エッジのインタフェースルーティングを行うことが想定される。

ノード ID の配布に関しては、各 Region で NICT の HANA (Hierarchical and Automatic. Number Allocation for locators)[15] をはじめとした自動的な Host ID 割り当て機構を配置することにより、容易に集約可能な Host ID の自動配布を実現できると考えている。

4.4 ノード内識別子

ノード内識別子について述べる。

本稿で述べた PPE のインタフェースは 400 Gbps×8 としていたが、データセンタや通信事業者によるインタフェースの論理的・物理的分割を考慮する場合及び、個々の企業等による個別使用等を考慮した際、低速な拡張 NIC 等を用いてインタフェース数の拡張を行うことを考慮する。例えば、Juniper 社の MX2020 ルータ[16] は、320 個の 100 GE インタフェースを 20 スロットに収容可能、10 GE インタフェースならば最大 1,920 個を 1 台に収容可能である。

このように、本稿で想定した PPE のインタフェース数 8 という数字は、最小スペックと言える。これについて本稿では、拡張 NIC の使用例を提示し、PPE の拡張性を考慮したうえで 16 bit をノード内識別子として与えた。

図 4 に拡張 NIC の使用による PPE の拡張性とノード内識別子の使用例を示す。

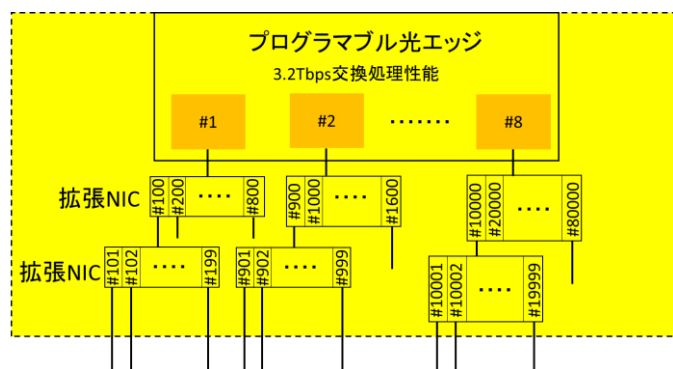


図 4. ノード内識別子の使用例

図 4 では、インタフェース #1 は拡張 NIC によりノード

内識別子が 800 個設定されており、またインタフェース #8 はノード内識別子 80,000 が設定されている例を示した。実際にどこまで拡張するかはユーザ次第となるが、ノード内識別子を 16 bit とすることで、キャリア毎、企業ごと、個別ユーザにまでインタフェースを割り当てることが可能となる。また SPN の実現を目指す 2020 年以降のネットワークの環境を想定すると、IoT デバイスの増加、ユーザの増加は十分考えられる。この課題に対しても、16 bit をノード内識別子として与えることで、上記のノード内識別子の割り当てをはじめとして、様々な用途に対応する柔軟なアドレスの割り当てが可能となる。

5. まとめ

本稿では、SPN で求められている、広域、マルチドメインに対応する新たなレイヤ 2 プロトコル、光 L2 において使用されるフレームにおいて、マルチドメインの一体運用、アドレス集約化によるルーティングの簡易化に焦点を当て、アドレス割り当てについて検討を行った。48 bit のアドレス空間において、先頭 2 bit を I/G, U/L 識別子、続く 10 bit を Domain ID, 8 bit を Region ID, 12 bit を Host ID, 16 bit をノード内識別子として与えた。以上のアドレス割り当てを行うことにより、国家に跨るネットワークの構築に加え、世界中の通信事業者レベルの分割や、国内における都市、都道府県レベルのアドレス集約化が可能となる。また企業ユーザから個別ユーザまでを対象とするインタフェースの割り当て等、使用者毎に各アドレスフィールドを柔軟に割り当てることが可能となる。以上より、2020 年以降における SPN での活用に対応するアドレス割り当てを行った。

6. 謝辞

本研究の一部は、NICT 委託研究「光トランスポート NW における用途・性能に適応した通信処理合成技術の研究開発」の成果です。

文献

- [1] K. Kitayama, A. Hiramatsu, M. Fukui, T. Tsuritani, N. Yamanaka, S. Okamoto, M. Jinno, and M. Koga, "Photonic Network Vision 2020 - Toward Smart Photonic Cloud," IEEE Journal of Lightwave Technology (JLT), Vol. 32, No. 16, pp. 2760-2770, August 2014.
- [2] 中尾彰宏, "新世代ネットワーク構想におけるネットワーク仮想化," 信学誌, vol.94, no.5, pp.385-390, May 2011.
- [3] S. Gringeri, N. Bitar and T. J. Xia, "Extending software defined network principles to include optical transport," in IEEE Communications Magazine, vol. 51, no. 3, pp. 32-40, March 2013.

- [4] 岡本聡, “スマートフォニックネットワークキング—アプリケーションドリブンネットワークキングの世界を目指して—,” 信学ソ大, BI-5-6, Sept.2015.
- [5] 岡本聡, 松本隼, 佐藤丈博, 山中直明, “光 L2 網実現を目指した仮想通信処理プロセッサによるプログラマブルノード構成,” 信学技報, Vol. 116, No. 205, PN2016-24, pp. 59-64, Sept. 2016.
- [6] 岡本聡, 鏝田幸大, 佐藤丈博, 山中直明, “光 L2 網における電気ラベル版フレーム構成,” 信学ソ大, B-12-7, Sept.2017.
- [7] 岡本聡, 菊田洗, 石井大介, 山中直明, “新世代広域レイヤ2網におけるアドレッシング及びエッジクロスコネクション設定手法の検討,” 信学技報, PN2010-15, Sept.2010.
- [8] 中島悠佑, 岡本聡, 山中直明, “光レイヤ2ネットワークにおけるスケーラブルなアドレス解決手法の提案,” 信学会 PN 研究会学生 WS, No.2, Mar. 2015.
- [9] K. Yarita, Y. Nakajima, A. Yamashita, J. Matsumoto, S. Okamoto and N. Yamanaka, "Cache retention time change method to reduce address resolution response time in wide area Layer 2 network," 2017 IEEE LANMAN, Osaka, 2017, pp. 1-2.
- [10] 総務省, “我が国のインターネットにおけるトラヒックの集計・試算,” 2017年5月
- [11] 総務省, “平成25年住宅・土地統計調査(速報集計)結果の要約”, 2013年10月
- [12] 経済産業省, “2017年版中小企業白書 概要”, 2017年4月
- [13] 人民日報, “中国の世帯数, 4億3000万で世界5分の1を占めトップ-人民網日本語版”, 2014年5月
- [14] 外務省, “国連加盟国一覧(アルファベット順)”, 2014年1月
- [15] “世界初, 広域ネットワークの自動構築に成功”, NICT NEWS, 2012年9号, pp.5-6, Sept. 2012.
- [16] MX2020 3D ユニバーサルエッジルーター
<https://www.juniper.net/jp/jp/products-services/routing/mx-series/mx2020/>