

セグメントルーティングにおけるセグメント階層化

三島 航¹ 明石 邦夫² 宇多 仁³ 篠田 陽一³

概要：本研究では、セグメントルーティングにおけるスケーラビリティ・安定性の向上のため、コントロールプレーンを階層化した階層的セグメントルーティングを提案する。階層的セグメントルーティングでは、上位セグメントの定義により下位セグメントの情報を隠蔽し、各ノードの持つ情報量を削減することでIGPの動作範囲を縮小する。本論文では、モデルの分類・比較を行い、データプレーンの拡張の有無による性能の差や、SR PCEの構成によるポリシー設定手法の違いについて議論を行った。得られた知見として、データプレーンの拡張を行わない場合は既存のデータプレーンを使用できるが、行わない場合に比べ、セグメントリストが増加する点、適用するネットワークの管理手法によりSR PCEのポリシー設定モデルを使い分ける必要がある点を得た。得られた知見を元に、今後の概念実証の実装に向けた方針を述べる。

キーワード：セグメントルーティング、トラフィックエンジニアリング、MPLS、NFV、SDN

Hierarchization of Segment Routing

Wataru MISHIMA¹ Kunio AKASHI² Satoshi UDA³ Yoichi SHINODA³

1. はじめに

携帯デバイスの普及や動画コンテンツの需要増に伴い、ネットワーク上のトラフィックは増加を続けている。大量のトラフィックが1つのノードに集中した場合、伝送遅延の増加や輻輳が発生するため、負荷に応じた経路変更が必要となる。また、ネットワーク機能を仮想化するNetwork Function Virtualization (NFV) やクラウドサービスの普及に伴い、ネットワーク上に提供されるサービス数が増加している。Deep Packet Inspection (DPI)・ファイアウォール・ロードバランサなどのサービスを適切に利用するため、特定の通信に必要なネットワーク機能へ通過させるサービスチェイニング [1] が提案され、実現のためサービス単位での経路制御が求められる。これらの要求に応じ、柔軟な経路変更によるトラフィック制御技術が提案されている。

セグメントルーティング (Segment Routing : SR) [2]

は、ネットワークをセグメントで表現し、セグメントを宛先や経路先として指定することで柔軟なトラフィック制御を実現する技術である。SRの課題として次の3つが挙げられる。1) SR domainに属する各ノードが全セグメントの識別子 (SID) の情報を保持する必要があるため、ネットワークの規模拡大に伴いテーブルが肥大化し、検索時間が増加する。2) パケットが宛先に到達するまでに経由するセグメントのSIDをセグメントリストとしてヘッダに挿入するため、複雑な経路を指定した際にヘッダサイズが増加する。3) ネットワークの規模が拡大するとセグメント情報の伝搬に利用するIGPの適用範囲・計算時間も増加し、安定性が低下する。これらの課題から、SRのスケーラビリティと安定性向上のため、各ノードの持つテーブル・セグメントリストの情報量とIGPの計算時間の削減が必要となる。本研究では、これらの課題を解決するための提案を行う。

2. 基盤技術と関連研究

2.1 SRの概要

SRの概念図を図1に示す。SRで用いられるセグメントには、ノードを表現するNode Segmentとノード間の特定のリンクを表現するAdjacency Segmentがある。SRに参加するノードの集合をSR domainと呼ぶ。SRでは各セグ

¹ 北陸先端科学技術大学院大学先端科学技術研究科
Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology

² 情報通信研究機構
National Institute of Information and Communications
Technology

³ 北陸先端科学技術大学院大学情報社会基盤研究センター
Reserch Center for Advanced Computing Infrastructure,
Japan Advanced Institute of Science and Technology

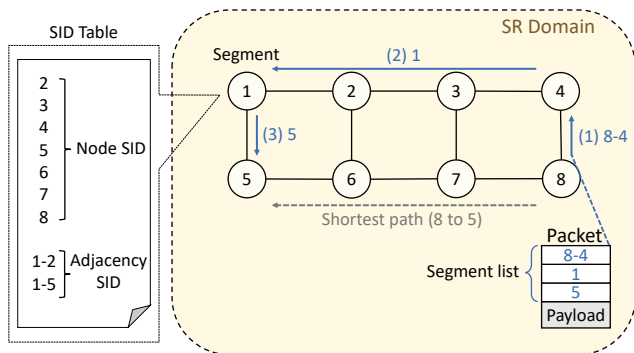


図 1 SR の概念

メントを Segment ID (SID) と呼ばれる識別子で識別する。SID には Node Segment を表す Node SID と Adjacency Segment を表す Adjacency SID がある。Node SID は SR domain 内で一意な Global Segment であり、Adjacency SID はノード内で一意な Local Segment である。

SR に参加するノードは SR domain 内の全セグメントの SID 情報を持つ。SR では SDN (Software Defined Network) への対応を目的とし、PCE (Path Computation Elements) による情報の管理 (SR PCE) が提案されている。

SR のコントロールプレーンには OSPF・IS-IS・BGP が、データプレーンには MPLS と IPv6 が提案されている。

SR では、SR domain での送信元となるノードが経路を指定する、ソースルーティングを採用している。経路情報は経由する全ての SID のリストであるセグメントリストをパケットに付加することで表現される。パケットを受信したノードは、セグメントリストの先頭を確認し、Node SID であった場合には経路表における宛先への送出インターフェースへ、Adjacency SID であった場合には該当リンクにつながるインターフェースへパケットを転送する。

図 1 では、SID 8 のノードが「8-4, 1, 5」というセグメントリストが付加されたパケットを受信した例を示している。SID 8 のノードから SID 5 のノードへの IGP での最短経路は、SID 7, SID 6 のノードを経由する経路である。(1) パケットを受信した際 SID 8 のノードは先頭セグメントを確認する。先頭セグメントは「8-4」という Adjacency SID であるため、SID 8 のノードは次の要素である「1」を先頭とし、8-4 が示すリンクがつながるインターフェースからパケットを送出する。(2) パケットを受け取った SID 4 のノードは、セグメントリストの先頭が Node SID を示す「1」であるため、セグメントリストを変更せず IGP による SID 1 が示すノードへの最短経路へパケットを送出する。SID 3・SID 2 のノードにおいても同じである。(3) SID 1 のノードは、セグメントリストの先頭が自らの Node SID である「1」であるため、次の先頭セグメントである「5」を確認する。そして SID 5 へパケットを送出する。SID 5 のノードは、セグメントリストの先頭が自らの Node SID で

ある「5」であるため、次の先頭要素を確認し、セグメントリストの最後の要素であるため SR による転送を終了する。

2.2 SR 拡張研究

SR におけるスケーラビリティの評価として、Sgambelluri らの例 [3] がある。この例では 2 つの実験が行われている。

1 つ目は、MPLS 網と光バイパスを持つマルチレイヤネットワークに SR を適用し、負荷に応じセグメントごとに光バイパスへの経路切り替えを行っている。この手法では SR PCE として SDN コントローラを用いており、経路切り替えのためのルールを OpenFlow[4] により定義し SR 網に展開している。検証で用いるネットワークでは、通常時は MPLS 網を利用し、消費帯域の量が事前に設定した閾値を超えた際に、SR PCE によりセグメントリストの構成と各ノードのフォワーディングテーブルを変更し、特定の通信経路を光バイパスへ動的に切り替える。この検証では、帯域を閾値とした動的経路制御においてパケットロスを起こさずに経路切り替えが可能であることが示されている。

2 つ目は、10×10 マンハッタンメッシュネットワークにおいて、異なるソースと宛先の組み合わせからなる 1 から 15 個のラベルを付加し、5000 回の試行を行うことで SR 動作検証と性能を計測する実験である。実験では SR で必要な経路計算・セグメントリストを表すラベルスタックの計算・ラベルプッシュなどの操作に要した時間を計測している。結果では 1 個のラベルでは 273.7 μ m, 15 個では 939.3 μ m を転送前の処理に要し、処理時間に 3 倍の差が生じたことと、転送中の処理時間は 1 個のラベルでは 50.6 μ m, 15 個では 53.1 μ m であり、与える影響が少ないことが示されている。この結果より、セグメントリストの増加が転送前のパス計算・ラベルスタック計算に大きく影響を与えることがわかった。

また、セグメントルーティングの経路制御の例として、Paolucci らの例 [5] がある。この例では、BGP FlowSpec でのサービスチェイニングを目的とし、経路制御を行うための手段として SR 網を採用している。この手法では、サービスを提供する 2 つのクライアントネットワークが 1 つのコアネットワークに接続したトポロジにおいて、BGP FlowSpec を拡張し、BGP Speaker と受信した Flow リクエストを処理する Flow Computation Element (FCE) の集合を SR PCE として利用することにより SR 網へポリシーを適用しサービスチェイニングを実現している。この手法の結果として、BGP FlowSpec の Flow メッセージを SR 網に拡張することで、BGP FlowSpec を通じた複数のネットワークに対する経路制御を実現可能なことが示された。

これらの先行研究は、負荷への適応やサービスチェイニングのための経路制御に SR を用いることを想定し行われている。先行研究から得られた知見として、Sgambelluri らの実験から、SR においてセグメントリストが増加する

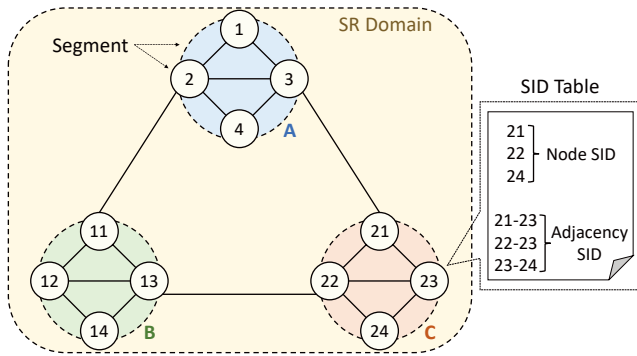


図 2 階層的セグメントルーティングの概念

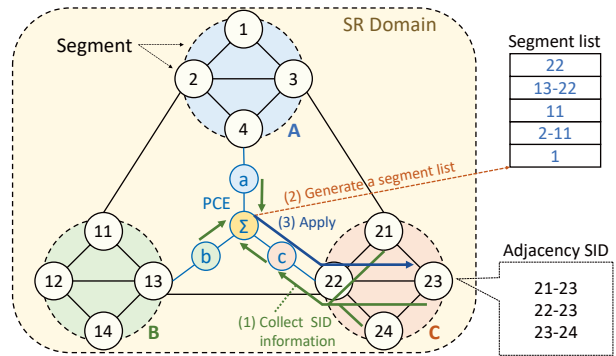


図 3 SR PCE による経路設定

と、経路計算やラベルスタックの計算時間が課題となるものの、転送時間は増加せず、スケーラビリティが確保されることがわかった。また、Paolucciらの手法により、複数のネットワークが接続したトポロジにおいても、ポリシーを適用し、SRによる経路制御が行えることがわかった。しかし、これらの手法では第1章で課題として挙げた各ノードの持つ情報量については扱っていない。また、Paolucciらの手法ではBGP FlowSpecを用いることを前提としており、使用がFlowSpecに対応したBGP環境に限られる。

これに対し、本研究ではSRを拡張することにより、各ノードの持つ情報量の削減・IGPの適用範囲の削減によるスケーラビリティと安定性の向上を目的として定める。

3. 提案

本研究ではSRにおけるスケーラビリティと安定性向上を実現するため、SRのコントロールプレーンに階層構造を導入した階層的セグメントルーティングを提案する。

3.1 階層的セグメントルーティングの概念と動作

階層的セグメントルーティングの概念図を図2に示す。

階層的セグメントルーティングでは、複数のセグメントを束ねた上位セグメントを定義することでSR domainを階層的に構成する。各下位セグメントは自分の属する上位セグメント内のSID情報のみを持つ。図2の例では、SID 23のノードが上位セグメントであるC内の「21, 22, 24」のNode SIDと、Adjacency SIDのみを所持している。他の上位セグメントに属するセグメントの情報を隠蔽することで、各セグメントの持つ情報を削減し、SRにおいてスケーラビリティの向上が可能となる。また階層的セグメントルーティングの利点として、コントロールプレーンとなるIGPの適用範囲を分割できる点が挙げられる。IGPの適用範囲を縮小することで、経路変更による影響範囲を狭め、経路再計算コストの削減や安定性の向上を実現できる。

図3, 図4に階層的セグメントルーティングの動作例を示す。以降の例では、SRの情報制御とポリシーに基づく経路計算にSR PCEを用いることを想定する。図3はPCE

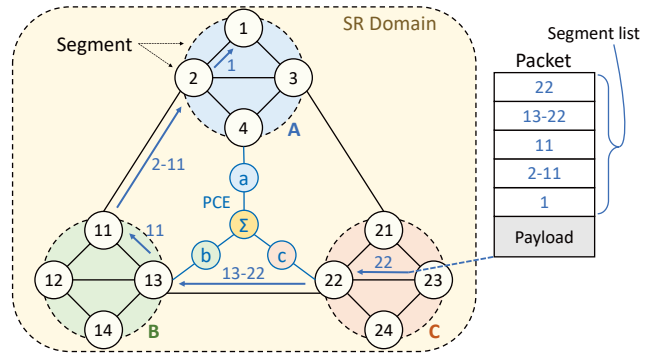


図 4 階層的セグメントルーティングによるパケット転送

がSR domain内のセグメント情報を収集し、セグメントリストを構築する様子を表している。SR PCE a, b, cが各上位セグメントA, B, Cを管理する下位SR PCEであり、SR PCE Σ がSR domain全体を管理する上位SR PCEである。(1)では、事前準備として各下位SR PCEが、それぞれの管理するセグメントの情報を収集する。その後、上位のSR PCEがそれぞれのSR PCEからSR domain全体のセグメント情報を収集する。(2)では、SR PCEが収集したセグメント情報を用い、事前に設定されたポリシーに基づいてセグメントリストを構築する。この例ではSID 23のノードからSID 1のノードへの経路を示すセグメントリストを構築している。(3)では、生成したセグメントリストを各ノードに展開する。この例では(2)で生成したセグメントリストをSID 23のノードに送信している。

階層的セグメントルーティングでは、各下位セグメントが持つ情報に合うようにセグメントリストを構築する必要がある。そのため、上位セグメントを隔てたNode SIDの指定は不可能であり、必ず上位セグメント間のAdjacency SIDを指定する形でセグメントリストを構築する。

生成されたセグメントリストを用いて実際にパケットを転送する様子を図4に示す。SID 23のノードがSID 1のノードを宛先としたパケットを送信する際、図3の処理で得たセグメントリストをパケットに付加し送信を行う。送信されたパケットは付加されたセグメントリストに基づき「22, 13-22, 11, 2-11, 1」の順にセグメントを経由し転送

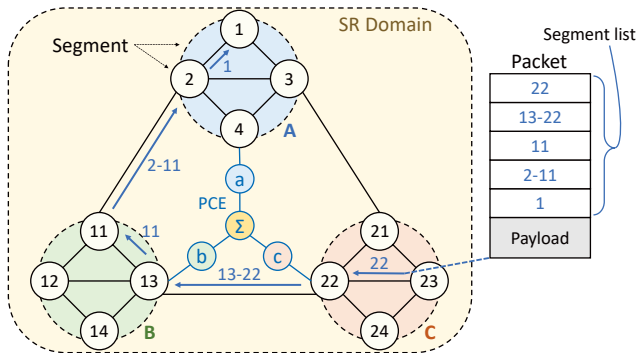


図 5 データプレーンの拡張が不要なモデル（展開モデル）

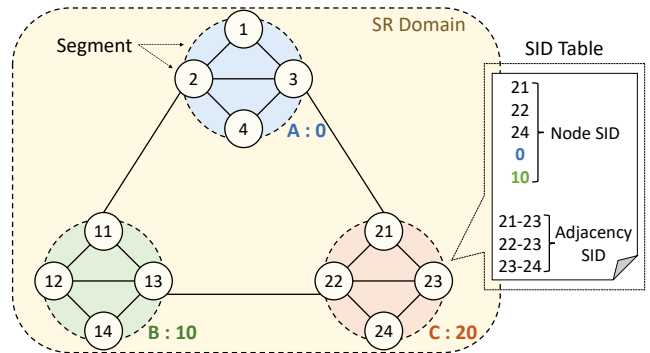


図 6 データプレーンの拡張が必要なモデル（上位 SID モデル）

が行われる。セグメントリストは各下位セグメントの持つ情報のみで構成されているため、特殊な動作をすることなく通常の SID の動作に基づいてパケットが送信される。これにより、SR domain の各ノードが持つテーブル量を削減しながら SR の動作が実現可能なことを示した。

3.2 設計モデルの分類

階層的セグメントルーティングはコントロールプレーン・データプレーンの違いにより複数のモデルが考えられる。本節では、階層型セグメントルーティングのモデルを以下の 2 つの軸で分類する。

- データプレーンへの変更の有無
- 階層化した SR PCE のポリシー設定

第 3.2.1 節、3.2.2 節でそれぞれの分類の特徴を示す。

3.2.1 データプレーンの拡張の有無による分類

提案手法では、コントロールプレーンの階層化を行い上位セグメントを新たに定義することで、各ノードの持つ情報量の削減を行なった。階層的セグメントルーティングの設計として、パケット転送における上位セグメントの扱いにより、データプレーンの拡張が不要なモデルと必要なモデルが存在する。以下にそれぞれの概念を示す。

3.2.1.1 データプレーンの拡張が不要なモデル

このモデルは、第 3.1 節で例として扱った手法である。これを展開モデルと呼ぶ。

図 5 に展開モデルの概念を示す。展開モデルでは各下位セグメントが自らの属する上位セグメント内部の情報のみを持つ。図 5 においても、SID 23 のノードが上位セグメントである C 内の「21, 22, 24」の Node SID と、Adjacency SID のみを所持しており、上位セグメント C の情報に閉じている。セグメントリスト構成時、SR PCE は各下位セグメントが持つ情報のみでパケット転送が行えるようにセグメントの展開を行う。構成されたセグメントリストには、上位セグメントを隔てたセグメントの指定は含まれず、上位セグメント間を移動する際は、必ず上位セグメント間の Adjacency SID が指定される特徴がある。

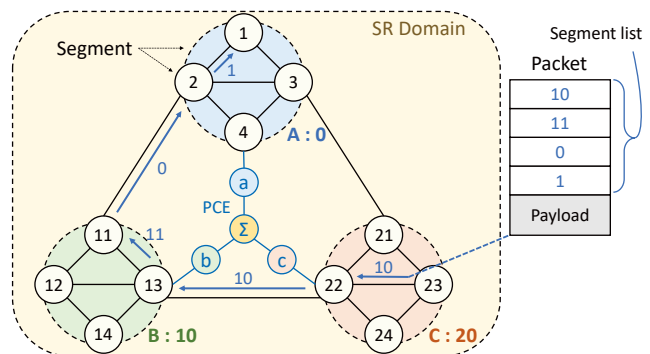


図 7 上位 SID モデルにおけるパケット転送

3.2.1.2 データプレーンの拡張が必要なモデル

このモデルでは、上位セグメントに割り当てた SID をパケット転送に用いる。これを上位 SID モデルと呼ぶ。図 6 に概念を示す。図 6 の通り、上位 SID モデルでは各上位セグメントに Global Segment である SID が割り振られる。また、各下位セグメントは自らの属する上位セグメント内部の情報に加え、他上位セグメントの SID の情報を持つ。

上位 SID モデルにおけるパケット転送の様子を図 7 に示す。上位 SID モデルでは、パケットの転送に上位 SID を使用できる。これにより上位セグメントを跨ぐ通信において、上位セグメント間の Adjacency SID ではなく上位 SID を直接指定することが可能となる。通常の SR では、Node SID は SR domain のノードと一対一で対応している。しかし上位セグメントモデルでは、複数のノードを同じ上位 SID で扱う。そのため、上位 SID モデルではある上位 SID が指定された際、宛先となるノード群からノードを選択しパケットを転送するように、データプレーンの動作を拡張する必要がある。また、上位 SID を登録するために SID テーブルの拡張も必要となる。

3.2.2 階層化した SR PCE のポリシー設定による分類

本論文では SR の情報管理と経路計算に SR PCE の利用を想定している。SR PCE は、対象となるセグメントの SID の情報を集め、ポリシーに従いセグメントリストを構成し、各ノードに配布する役割を持つ。SR PCE の設計として、上位 SR PCE へポリシーを設定し下位の SR PCE

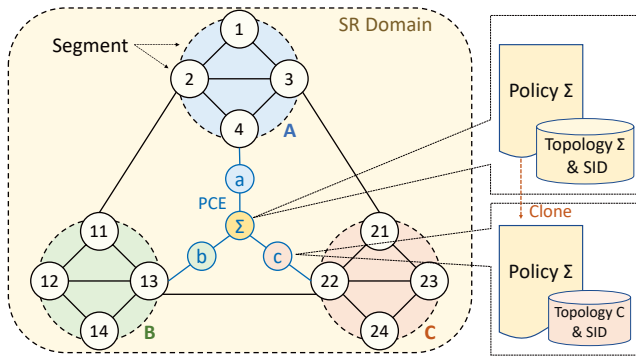


図 8 配布モデル

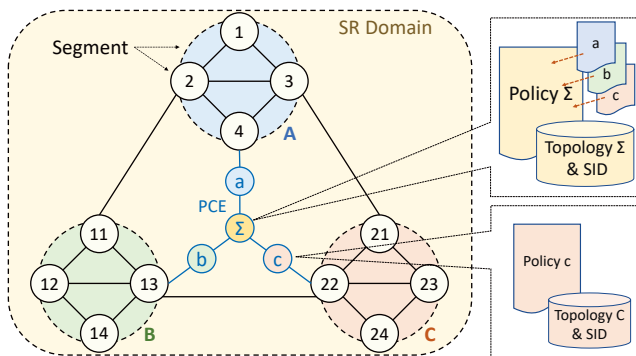


図 9 収集モデル

へとポリシーを配布するモデルと、下位の SR PCE へポリシーを設定し上位の SR PCE が収集し管理するモデルの二つが考えられる。以下にそれぞれの概念を示す。

3.2.2.1 上位 SR PCE のポリシーを配布するモデル

上位 SR PCE に設定したポリシーを下位 SR PCE に配布するモデルを配布モデルと呼ぶ。図 8 に概念を示す。

配布モデルでは、ネットワークの管理者が上位の SR PCE にのみポリシーを設定し、上位 PCE が必要な設定のみを下位の各 SR PCE へと配布することで、SR ドメイン全体にポリシーを設定する。配布モデルの特徴として、ポリシーの一元管理が可能な点や、上位・下位の PCE を一体として扱いやすい点が挙げられる。

3.2.2.2 下位の SR PCE のポリシーを収集するモデル

下位の SR PCE に設定されたポリシーを上位 SE PCE が収集するモデルを収集モデルと呼ぶ。図 9 に概念を示す。

収集モデルでは、それぞれの上位セグメントにおいて各管理者が下位 PCE にポリシーを設定し、上位 PCE が下位 PCE からポリシーを収集することで SR domain 全体のポリシーを作成する。図 9 に、下位 SR PCE である a, b, c のポリシーを収集し、SR domain 全体のポリシーを Σ に構築した様子を示した。収集モデルの特徴として、上位セグメント毎のポリシーの分離が容易であり、異なるポリシーで運用されるネットワーク間の連携に適す点が挙げられる。

4. 議論

4.1 階層的セグメントルーティングの利点・問題点

本研究では SR におけるスケーラビリティと安定性の向上のため、SR のコントロールプレーンを階層化した階層的セグメントルーティングを提案した。階層的セグメントルーティングの利点として、各ノードの持つ情報量削減・IGP のエリア分割・セグメントリスト長の削減が挙げられる。階層的セグメントルーティングの問題点としては、階層化した PCE 間の問い合わせ時間や、上位セグメントを超えた Node Segment の指定ができないことによるセグメントリスト長の増加が挙げられる。

PCE 間の問い合わせ時間の発生は、上位 PCE と下位 PCE の距離に起因するものであるため、ネットワークの規模に関わらず一定のコストとなることが予想される。そのため、階層化による他の利点との比較を、ネットワークの規模毎に行うことで評価を行う。上位セグメントを超えた Node Segment の指定については、第 3.2.1.2 節で触れた上位 SID モデルを採用することで解決が可能である。また、その他のモデルについても、セグメントリスト長は増加するもののパケット転送は実行できるため、階層化による情報の隠蔽の利点、セグメントリスト長が増加する問題点の比較を行いながら研究を進める。

4.2 階層的セグメントルーティングの設計手法の比較

第 3 節で、階層的セグメントルーティングのモデルを 2 つの軸に基づいて分類した。この節ではそれぞれの利点・問題点の比較を行う。データプレーンの拡張の有無により、上位セグメントに付与した SID を用いパケット転送を行う上位 SID モデルと、データプレーンは拡張せず、他上位セグメントに属するノードの情報を隠蔽したままパケット転送を行う展開モデルを分類した。上位 SID モデルでは、上位セグメントを超えたパケット転送において上位 SID を Node SID として指定可能である点や、他の上位セグメントを Node SID で扱う性質上、他上位セグメント内でトポロジ変更が生じた際も、それによるセグメントリストの変更が生じないという利点がある。但し、上位 SID モデルではデータプレーンを拡張し、複数のノードを同一の上位 SID で扱うための仕組みが必要となる。第 4.3 節で具体的な手法を議論する。それに対し展開モデルではコントロールプレーンのみの拡張で実現可能であり、セグメントリストは各上位セグメント内の情報のみで扱えるように展開される。但し、展開モデルでは上位セグメントを超えた Node SID の利用が不可能であるため、上位セグメントを超える際は必ず上位セグメント内のエッジノードへ転送した後上位セグメント間の Adjacency SID で転送を行う必要があり、上位セグメントモデルに比べセグメントリストが

増加する。また、上位セグメント間に等コストのリンクが複数存在する場合、ロードバランスのためにノードで複数のセグメントリストを記憶し、使い分ける必要がある。

階層化した SR PCE のポリシー構成により、上位 PCE にポリシーを設定し、下位 PCE へと配布する配布モデルと、それぞれの下位 PCE にポリシーを設定し、上位 PCE が集約を行う収集モデルを分類した。配布モデルではポリシーの一元管理が可能であり、また上位・下位の全ての SR PCE を一体として扱いやすい利点が存在する。収集モデルではポリシー設定者の分離が容易であり、異なるポリシーで運用されるネットワーク間の連携に適している。収集モデルでは各管理者のポリシーを矛盾なく統合し、各下位セグメントへと配布するための機構が必要となる。管理者が同一であり、物理的な位置によって上位セグメントを分割するようなネットワークでは、配布モデルによるポリシーの一元管理が優位である。一方で、管理者の異なる上位セグメントを自律分散的に構成しそれらを SR によって接続するようなネットワークでは、収集モデルにより各管理者のポリシーを適用することが求められる。このように、適用するネットワークの種類によりモデルの使い分けが必要となる。

4.3 実装計画

第3章では階層的セグメントルーティングの概要と分類を示し、評価を行った。評価の結果、本研究では実装の簡易さを考えて、まず展開モデルの実装を行う。PCE の設計については、上位・下位の SR PCE を一体として扱える利点から、配布モデルでの実装を行う。概念実証として階層的セグメントルーティングを実装し、試験環境において動作させ、適切な上位セグメントサイズ・問い合わせ時間とテーブルサイズのトレードオフについて評価を行う。動作の検証として、SR においてセグメントリストの要素数による処理時間を計測し、何段までのスタックを行うべきかを議論し、セグメントリストを削減可能である階層的セグメントルーティングのスケラビリティについて優位性を検証する。また、ノードが定期的に出現・消滅するネットワークにおいて、通常の SR と階層的セグメントルーティングの動作を比較し、安定性を計測する。

展開モデルの実装を行った後、上位 SID モデルの実装も行い比較する。上位 SID の実装方法として、SRv6 においてエニーキャストアドレスの使用を検討している。

5. おわりに

本研究では、セグメントルーティングのコントロールプレーンを階層化した階層的セグメントルーティングを提案した。階層的セグメントルーティングでは、小さなセグメントの集合を扱う上位セグメントを定義し、他のセグメントの情報を隠蔽する。これにより SR のスケラビリティ

と安定性の向上が可能となる。この階層型セグメントルーティングを実現するために、設計モデルを上位セグメントの扱いと SR PCE でのポリシー設定手法という二つの軸で分類を行なった。分類の結果、上位セグメントの扱いでは、通常の SR のデータプレーンでパケット転送を行う展開モデルと、データプレーンも拡張し、上位セグメントを表す上位 SID を宛先としたパケット転送を実現する上位 SID モデルに分類した。また、SR PCE でのポリシー設定手法では、上位 SR PCE でポリシーを一元管理する配布モデルと、各下位 SR PCE にポリシーを設定し、上位 SR PCE が収集することで SR domain 全体のポリシーを構成する収集モデルに分類した。

得られた知見として、上位 SID モデルではセグメントリストの削減が可能であるが、データプレーンの拡張が必要となること、展開モデルでは既存のデータプレーンを使用可能であるが、上位セグメントを超えた転送において、上位 SID モデルに比べセグメントリストが増加することが得られた。また、SR を適用するネットワークの管理形態により、SR PCE のポリシー設定モデルを使い分けるべきであることが得られた。

今後は概念実証としてコントロールプレーンに展開モデルを、PCE に配布モデルを実装し動作の検証を行う。またその後上位 PCE モデルについても実装を行い、展開モデルとの比較を行う。

謝辞

本研究を進め、論文を執筆するにあたり、情報通信研究機構北陸 StarBED 技術センターの井上 朋哉博士に多大なご助力を頂きました。心より感謝いたします。

参考文献

- [1] Joel Halpern and Carlos Pignataro. Service function chaining (sfc) architecture. Technical report, 2015.
- [2] Clarence Filisfilis, Nagendra Kumar Nainar, Carlos Pignataro, Juan Camilo Cardona, and Pierre Francois. The segment routing architecture. In *Global Communications Conference (GLOBECOM), 2015 IEEE*, pages 1–6. IEEE, 2015.
- [3] Andrea Sgambelluri, Alessio Giorgetti, Filippo Cugini, Gianmarco Bruno, Francesco Lazzeri, and Piero Castoldi. First demonstration of sdn-based segment routing in multi-layer networks. In *Optical Fiber Communications Conference and Exhibition (OFC), 2015*, pages 1–3. IEEE, 2015.
- [4] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. Openflow: enabling innovation in campus networks. *ACM SIGCOMM Computer Communication Review*, 38(2):69–74, 2008.
- [5] Francesco Paolucci, Alessio Giorgetti, Filippo Cugini, and Piero Castoldi. Service chaining in multi-layer networks using segment routing and extended bgp flowspec. In *Optical Fiber Communication Conference*, pages W4J–3. Optical Society of America, 2017.