修士論文 2019 年度 (令和元年)

IPv6 シングルスタックネットワークにおける ダイナミックなアドレス変換テーブル広告手法

慶應義塾大学大学院 政策・メディア研究科
豊田安信

修士論文要旨 - 2019 年度 (令和元年度)

IPv6 シングルスタックネットワークにおける ダイナミックなアドレス変換テーブル広告手法

2019 年現在，IANA が保有する IPv4 アドレスプールは既に枯渇しており，各地域レジストリからも 2020 年頃には新規割当が行えなくなることが予想されている．一般に新規に IPv4 アドレスの取得するためにはこのような民間取引市場を利用する方法が考えられるが，1 アドレスあたりの単価は年々上昇しており，新規に IPv4 ネットワークを構築するためのコストは日々上昇していくことが考えられる．IDC 事業者・コンテンツ事業者がビジネスを拡大するためには，IPv4 アドレスを極力使用しない IPv6 シングルスタックネットワークの活用が不可欠になっている．

一方で 2019 年現在においても IPv4 によるアクセス・トラフィックは依然としてインターネット全体の大きな割合を占めていることから，IPv6 シングルスタックネットワークでありながら IPv4 によるサービスを継続して提供可能なネットワーク設計が必要になってくると言える．

IPv6 のみ構築された IPv6 シングルスタックネットワークにおいて既存の IPv4 クライアントに対してサービスを提供する方法として、ステートレスアドレス変換を利用した”SIIT-DC” と呼ばれるネットワークデザインがインターネット標準として標準化されている．SIIT-DC では BR(Border Relay) と呼ばれる変換ノードを IPv4 ネットワーク・インターネットとの境界点ごとに設置し，明示的アドレス変換テーブル (EAMT: Eplicit Address Mapping Table) を参照してプロトコル変換を行い，IPv6 ノードでの IPv4 サービス提供を可能にする．しかしながら SIIT-DC では EAMT の動的な交換方法についての定義がなされておらず，対外接続点が複数存在する場合の冗長性の維持が難しい点や，IPv4 でサービス提供を行なうサーバーの構成変更が行われた場合に運用負荷が非常に高くなる点が課題に挙げられる．

本研究では BGP を利用したアドレス変換テーブルの広告・更新技術と，それを適切に運用するために必要なノード群の設計手法を提案する．これにより，SIIT-DC の課題であった冗長性の維持や構成変更へのに対して，ダイナミックに対応することが可能になる．

この手法を評価するために，新たに BGP によるアドレス変換テーブル制御機構を実装したソフトウェアルーターを実装し，多くの対外接続点を持つ学術 ISP である WIDE Project のバックボーンネットワークをモデルケースに，エミュレータを用いて概念検証実験を行った．考えられる他の手法と比較し，本手法が冗長性と柔軟性の点で優位であることが証明された．

キーワード:
1. IPv6, 2. データセンターネットワーク, 3. ネットワークオペレーション, 4. IPv6 移行技術

慶應義塾大学大学院 政策・メディア研究科
豊田安信

Abstract of Master's Thesis - Academic Year 2019

Dynamic advertising method of
Explicit Address Mapping in IPv6 single stack network.

Dynamic advertising method of Explicit Address Mapping in IPv6 single stack network.
Keywords :
1. IPv6, 2. Data center network, 3. Network operation, 4. IPv6 transition mechanism

Keio University Graduate School of Media and Governance
Yasunobu Toyota

目 次

図 目 次

表 目 次

第1章　序論

本章では本研究の背景とモチベーション，および全体の構成について記述する．

1.1　IPv6 シングルスタックネットワークに求められる役割

1.1.1　IDC ネットワークを取り巻く環境

IDC 市場の広がり

近年，ライブ映像配信のようなリアルタイムなサービスに対するニーズが年々高まっている．例えば Cisco 社の調査 [2] によれば，2022 年には全てのアプリケーショントラフィックのうちインターネットビデオが有する割合が 82 %を超え，そのうち 17 %がライブ映像配信が占めると予想されている．リアルタイムな高品質サービスを提供するためには，ユーザーの地理的に近いサービス拠点から配信を行うことが有効であるため，今後 IDC・コンテンツ事業者が各地域拠点を介したコンテンツ配信基盤を活用するしていくことが予想される．

一方で，インフラストラクチャに対する災害や地政学的リスクの軽減は，コンテンツ事業者の継続的な事業の成長のためには避けては通れない課題である [3]．2011 年に発生した東日本大震災以降，国内の IDC 事業者やコンテンツ事業者を中心に，関東大都市圏に集中していたサービス拠点への依存性を解消するために，東京圏以外の各地域にサービス拠点を分散する取り組みが活発だ [4]．大阪・名古屋の他の都市圏の IDC は 2019 年現在満床状態が続いているほか，他の地方拠点都市も含めた IDC 建設も並行して行われている．

特に近年では VXLAN や SRv6 のような新しいネットワーク仮想化技術の標準化も進み，サービス拠点のマルチテナンシーと柔軟性を両立するネットワークデザインの障壁が低くなってきているため，今後より多くの IDC・コンテンツ事業者のサービス拠点の拡大が続くと想定できる．

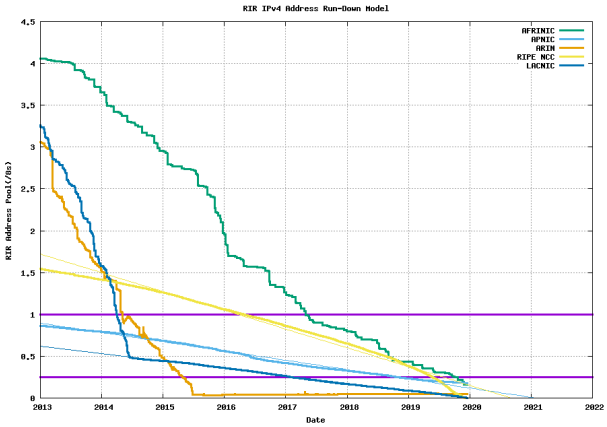

図 1.1: Projection of consumption of Remaining RIR Address Pools. potaroo.net より引用 [1]

IPv4 アドレスの枯渇

2019 年現在，IANA[1]が保有する IPv4 アドレスプールは既に枯渇しており [5]，各 RIR[2] からも 2021 年頃までには新規割当が行えなくなることが予想されている [1].

一方で近年は民間事業者間アドレス取引も盛んに行われている．一般に新規に IPv4 アドレスの割当を受けるためにはこのような民間取引市場を利用する方法が考えられるが，1 アドレスあたりの単価は年々上昇傾向にあり [6]，新規に IPv4 ネットワークを構築するためのコストは日々上昇していくことが考えられる．

IPv4/IPv6 デュアルスタックネットワークの問題

IPv6 プロトコルの導入に主に用いられていた手法として IPv4/IPv6 デュアルスタックネットワークが挙げられる [7]．IPv4/IPv6 デュアルスタックネットワークとは，IPv4 ネットワークと IPv6 ネットワークを同一機器群上に並行して運用する手法であり，企業・一般家庭向けアクセスネットワークを中心に IPv4/IPv6 デュアルスタック環境の整備が進んでいる．

一方でコンテンツ事業者が運用する IDC では以下の主な 3 つの理由からデュアルスタック環境の導入はデメリットが大きい．

- **IPv4 アドレスの継続的調達が困難**
  先に述べたように，IPv4 アドレスをサービスの成長にあわせて継続的に調達していくことは困難である．民間市場の市況に調達コストが左右されるため長期的な見通しが立てにくい．

[1]Internet Assigned Numbers Authority．インターネットに利用される様々な資源を一元的に管理する組織．https://www.iana.org/
[2]Regional Internet Registry.

- **オペレーションコストの肥大化**
  デュアルスタック環境では 2 つの異なる IP プロトコルを同時に運用する必要があるため、シングルスタック環境と比べて運用コストの上昇が見込まれる [?].
- **ネットワーク機器の性能要件の上昇**
  デュアルスタック環境では，シングルスタック環境よりも多くの経路をネットワーク機器が保持しなければならないため，より高性能な機器を導入する必要がある．

1.1.2 IPv6 シングルスタックネットワーク

IDC 事業者・コンテンツ事業者がビジネスを健全に拡大するためには，IPv6 ネットワークのみで機器間を接続した IPv6 シングルスタックネットワークの利用が不可欠である．

IDC の IPv6 シングルスタックネットワークには以下のような働きが期待される．

IPv4 サービスの提供

Google 社が定常的に行っている調査 [8] によれば，2019 年 12 月現在全世界のインターネットトラフィックの 7 割程度を IPv4 トラフィックが依然として占めている．将来的には IPv6 によるアクセスの割合が徐々に大きくなることが予想されるが，今後しばらくは IPv4 クライアントに対しても IPv6 クライアントと同等にサービス提供を行っていくことが望ましい．

コンテンツ事業者の IPv6 シングルスタックネットワークにおいても，何らかの手段を用いて IPv4 サービスを継続して提供する機構を備える必要がある．

シングルスタック運用による OPEX/CAPEX の削減

第 1.1.1 項で述べたように，IPv4/IPv6 デュアルスタックネットワークではオペレーションコストの肥大化が問題視されていた．IPv6 シングルスタックネットワークでは IPv4 ネットワークを廃止することが出来るため，OPEX[3] と CAPEX[4] の軽減が期待される．また IPv6 アドレスは IPv4 アドレスと比較して広大なアドレススペースを有するため，アドレススペースに依存しない柔軟なネットワーク設計が可能になる．

IPv4/IPv6 デュアルスタックネットワークと同等以上の性能

IPv6 により提供されるサービスはもちろんのこと，IPv4 によるサービスにおいても IPv4/IPv6 デュアルスタックネットワークと同等の耐障害性・サービス品質・サービス容量が保証されることが望ましい．

とりわけネイティブな IPv4 ネットワーク以外の手段を用いて提供される IPv4 サービスの性能の担保が運用課題になると予想される．

[3] Operating expense. 運用に掛かる継続的なコスト.
[4] Capital expenditure. 設備配備に掛かる初期投資コスト.

1.2 本研究のモチベーションと取り組み

第 1.1.2 項で述べたような IPv6 シングルスタックネットワークに求められる要件のうち，IPv4 サービスの提供における冗長性や構成変更への追従性の向上を促す手法の確立を目指す．

本研究では IPv6 シングルスタックネットワークにおける IPv4 サービスの提供手法のうち，アーキテクチャがシンプルで広範な利活用が期待される SIIT-DC[9] に着目した．SIIT-DC とは IPv6 ネットワークと IPv4 ネットワークの各境界部に，BR[5]を配備することにより，IPv6 ネットワークのみに属するホストで仮想的に IPv4 サービスを提供するネットワーク設計を定めたインターネット標準である．SIIT-DC において各 BR は静的に定義されたアドレス変換テーブルを利用してネットワークプロトコル変換を行うため，BR を複数配備する場合における一貫性の確保や冗長性，IDC 内の構成変更に対する追従性の面で課題があった．

本研究では動的経路アルゴリズムの一つである BGP[11] を利用したメッセージングによるアドレス変換テーブルの動的な広告手法を提案する．エミュレータを利用した概念実証実験により，本提案手法がこれらの課題に対して効果的に作用することが証明された．

1.3 本論文の構成

本論文の構成を以下に示す．

第 2 章では，IPv6 シングルスタックネットワークにおける IPv4 サービス提供手法に関してそれぞれの特徴や利点を紹介し比較する．

第 3 章では，IPv4/IPv6 プロトコル変換を利用した IPv4 サービス提供手法の一つである SIIT-DC のアーキテクチャと，解決すべき課題について述べる．

第 4 章では，SIIT-DC の課題を解決するために考えられる手法を比較・検討する．

第 5 章では，本研究において提案するダイナミックなアドレス変換テーブル広告手法の要件と構成について記述する．またメッセージングプロトコルとして採用した BGP の技術的利点について述べる．

第 6 章では，本提案手法の BGP メッセージペイロードの設計と第 7 章でも評価実験に用いる PoC の具体的な実装について紹介する．

第 7 章では，第 3 章で述べた課題に対して，本提案手法が有用であることを検証するための実証実験の概要及び具体的なシナリオについて述べ，結果を考察する．

第 8 章では，本研究のまとめと今後のロードマップについて検討する．

[5]Border Relay. IP/ICMP 変換アルゴリズム [10] を実装した IPv4/IPv6 トランスレーション機器．

第2章　IPv6シングルスタックネットワークでのIPv4サービス提供手法

本章ではIPv6シングルスタックネットワークでのIPv4サービス提供手法を比較し，検討する．

2.1　概要

第1.1.2で述べたように，コンテンツ事業者が運用するIPv6シングルスタックネットワークの重要な役割の一つに，IPv4クライアント端末に対するサービス提供がある．

関連して，アクセスネットワーク網ではIPv6シングルスタックネットワーク上でIPv4によるインターネット接続をクライアントエッジに提供する手法はをIPv4aaS[1]と呼称し，様々な手法が検討されている[12]．

一方でコンテンツ事業者が運用するネットワークでのIPv4サービス提供においては下のような要件を満たす必要があるため，必ずしもアクセスネットワークでのIPv4aaSと同様の方法が適切であるとは限らない．

2.1.1　IPv4サービス提供機構に求められる要件

IPv4クライアントからのアクセス

IPv4クライアントに対して透過的にサービスを提供する機構を備える．一般的なサーバークライアントモデルを想定した場合，インターネット上のIPv4クライアントからサービス提供サーバーに能動的に接続するためには，FQDN[2]もしくはIPv4アドレスをIPv4クライアントが指定出来る必要がある．

[1] IPv4 as a Service

[2] Fully Qualified Domain Name. 完全就職ドメイン名

スケーラビリティ

近年のコンテンツ事業者のネットワークでは，サービスのニーズに合わせて柔軟にスケールアウト[3]可能な設計であることが重要視されている [13]．同様に IPv4 サービスの提供手法に関しても，事業者の IPv4 サービス規模の変化にあわせて柔軟に拡大・縮小可能なアーキテクチャが求められる．

例えば，第 1.1.2 でも述べたように，将来的に IPv4 クライアントの占める割合が IPv6 クライアントに相対して低下していった場合に，既設の IPv6 ネットワークへの影響を最小限にしつつ，IPv4 サービス提供機構を縮小可能であると望ましい．

IPv4 ネットワークへの非依存性

第 1.1.2 項で述べたように，IPv6 シングルスタックネットワークのメリットを最大限に活かすためには IPv4 サービスを提供する場合においても IPv4 ネットワーク及びアドレスに極力依存しないことが望ましい．

2.2 IPv4 サービス提供手法の分類

想定される IPv4 サービス提供機構をその技術的差異や狙いを基に以下の 3 つの手法に分類した．

2.2.1 L7 リバースプロキシ

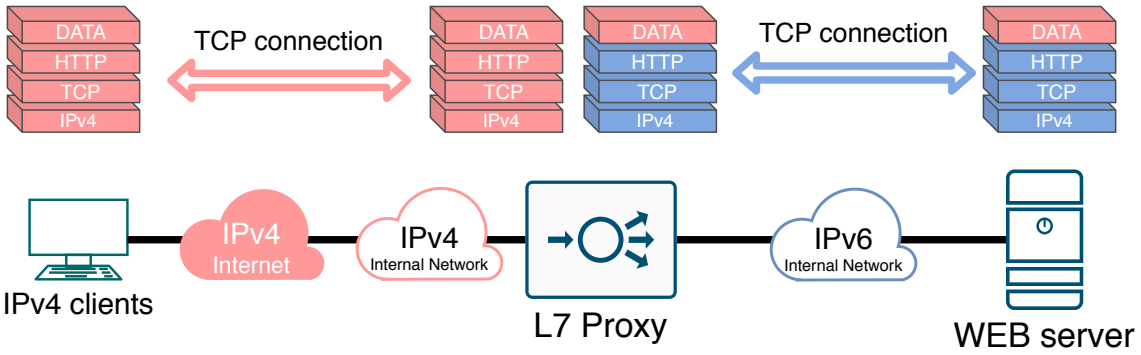

図 2.1: L7 リバースプロキシによる IPv4 サービス提供

L7 リバースプロキシとは，クライアントからの接続をプロキシサーバーがアプリケーション層レベルで終端し，プロキシサーバーがクライアントに代わってサーバーと接続す

[3]水平スケール．同等性能の機器を増減させることでサービス容量を拡大・縮小可能なモデル．

る機構である [14]．図 2.1 に本手法の構成を簡便に示す．主に WEB サーバーへの HTTP 接続を負荷分散するための手法として広く採用されている．

IPv6 シングルスタックネットワークにおいて IPv4 サービスを提供するためには，IPv4 インターネットとの接続点からプロキシサーバーまでの間に IPv4 ネットワークを配備する必要がある．

IPv4・IPv6 間のプロトコル仕様の差を考慮する必要が無いため互換性に留意する必要が無い点や，MTU[4]を減らさずにアプリケーショントラフィックを伝送可能である点が利点に挙げられる．

一方でアプリケーションレイヤーでのコネクション終端やそのステート管理を行う必要があるため，プロキシサーバーに負荷が掛かるため高性能な機器の導入が必要になる．

またスケールアウトを可能にするために L4LB と組み合わせた 3 ステージのアーキテクチャ利用する手法が近年主流である [15, 16] が，この手法を採用するためには，ある程度の IPv4 ネットワークを配備する必要があり，第 2.1.1 項で述べた要件に合致せず，IPv6 シングルスタックネットワークのメリットを損なうことになる．

2.2.2 IPv4/IPv6 トンネリング

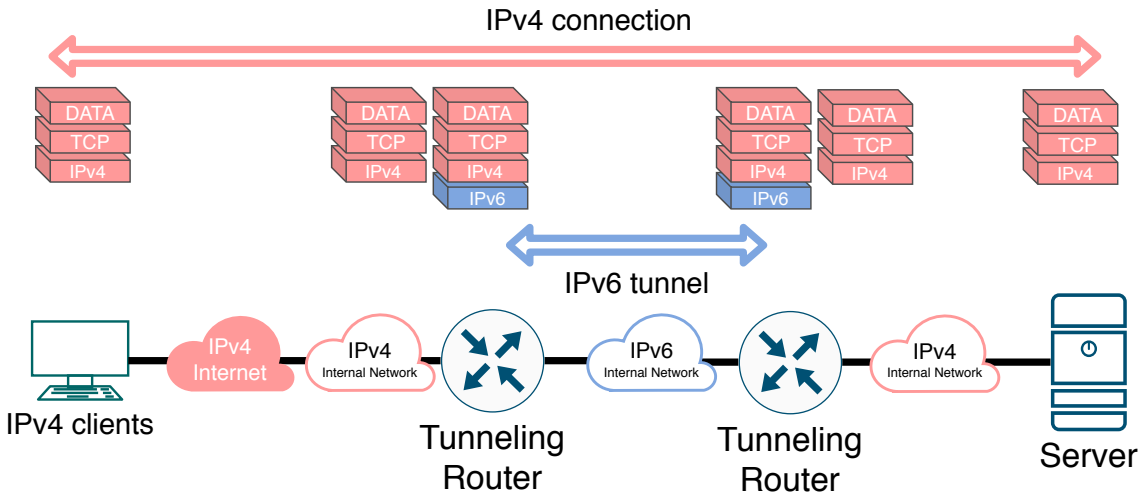

図 2.2: IPv4/IPv6 トンネリングによる IPv4 サービス提供

IPv4/IPv6 トンネリングとは，IPv4 パケットを IPv6 パケットによってカプセリングすることで IPv6 ネットワークを通過させる手法である．IPv4 トラフィックを透過的に利用することが出来るため，アクセスネットワークで最も一般的に利用されている IPv4aaS 手法である [12]．図 2.3 に本提供手法の構成を簡便に示す．

IPv6 シングルスタックにおける IPv4 サービス提供手法としては，IPv4 クライアントから到達したパケットをトンネルルーターによって一度 IPv6 パケットでカプセリングし，

[4]Maximum Transmission Unit. ここでは一つのパケットに搭載可能なデータ量を指す．

IDC 内の IPv6 シングルスタックネットワークを通過させ，IPv4 サービス提供サーバー上もしくはその直前で再びでカプセリングを解くことで，IPv4 提供サーバーまでネイティブな IPv4 トラフィックを通過させる運用が考えらえる．IPv4 ネットワークをサーバーで利用できるため，多種多様なアプリケーションでの採用が期待できる．

しかしながら，トンネルルーターと IPv4 サービスサーバー間にある程度の IPv4 ネットワークを配備しなければならず，ToR[5]及びサーバーでは IPv4/IPV6 デュアルスタック運用が必要になるため，第 2.1.1 項で上げた要件である「IPv4 ネットワークへの非依存性」に合致しない．また，トンネルプロトコルの多く [17] はは基本的に 1:1 もしくは 1:N の接続が基本となるため，水平スケールさせることが困難である．

2.2.3 IPv4/IPv6 トランスレーション

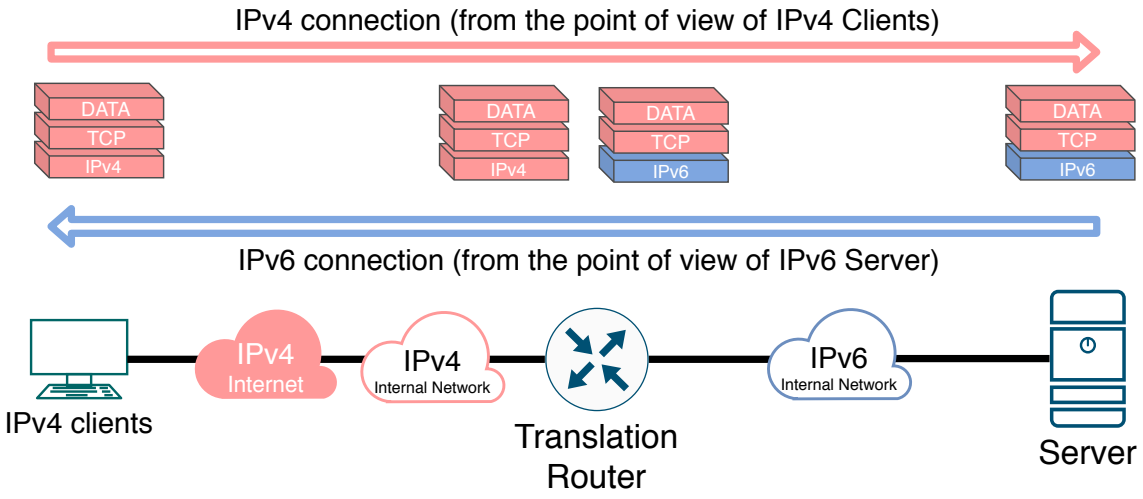

図 2.3: IPv4/IPv6 トランスレーションによる IPv4 サービス提供

IPv4/IPv6 トランスレーションとは，IPv4 パケットと IPv6 パケットを IP/ICMP 変換アルゴリズムを利用して相互に変換する手法である．$1:N$ の関係でアドレス・ポート変換を行うステートフルな NAT64[18] と，1 : 1 でアドレス変換を行うステートレスな SIIT[10] が定義されている．IPv4 ネットワークと IPv6 ネットワークの境界に位置する変換ルーターにより，相互にプロトコル変換が行われる．

IPv4/IPv6 トランスレーションでは IPv4 アドレスを IPv6 アドレスとして表現することが要求されるが，変換プレフィックスと呼ばれる IPv6 ネットワークプレフィックスに IPv4 アドレスを埋め込むことで，任意の IPv4 アドレスを IPv6 ホストから認識可能な形で表現する．変換プレフィックスには RFC6052 で定義された 64:ff9b::/96 の他に，運用者が専有可能な GUA[6]の/96 の IPv6 プレフィックスを利用することが想定されている [19]．

[5]Top of rack switch．ここではサーバーの L2 終端を行うルーターを指す．
[6]Global Unicast Address.

図 2.3 で示すように，変換ルーター以外のホストが IPv4 ネットワークに属する必要が無いため，第 2.1.1 項で述べた「IPv4 ネットワークへの非依存性」の面で，他の 2 手法より優れていると言える．また，IPv4 サービスを行うサーバーから変換ルーターの間はネイティブな IPv6 ネットワークで接続可能なため，ECMP[20] による経路の冗長化が可能なほか，ステートレスモードでは変換ルーターの水平スケールが可能な点で，IPv6 シングルスタックネットワークにおける IPv4 サービス提供に求められる要件を満たしやすい．

一方で IPv4 と IPv6 のプロトコル実装に差があるため，コンテンツ事業者のサービスの内容によってはサービス影響を考慮する必要がある点は留意すべきである．

第3章　SIIT-DCのデザインと現状の課題

第2.2.3で述べたPv4/IPv6トランスレーションを用いたIPv4サービス提供手法の一つとして，SIIT-DCがインターネット標準化されている．本章ではてSIIT-DCのデザインとメリット及び考えられる運用，そして現状の課題について述べる．

3.1　SIIT-DC

3.1.1　概要

SIIT-DCとは，ステートレスIP/ICMP変換アルゴリズム[10]を利用して，IPv4インターネット・ネットワークからのアクセスをIPv6シングルスタックネットワーク上のホストに提供するためのネットワークデザインである．2016年よりIETF IPv6 Operations WG[1]によりインターネット標準化(Informational RFC)されている[9].

3.1.2　用語

SIIT-DCで利用される用語及び特殊な役割を有する機器・技術について述べる．

SIIT

SIIT[2]とはIPv4/IPv6トランスレーションに用いられるプロトコル変換機能の略称である．RFC2765[21]で初めて標準化され，その後RFC6145[22]により一部の仕様が実運用のユースケースに合わせて変更され，現在はIPv6拡張ヘッダーを扱う機構などが追加されたRFC7915[10]が現行の標準仕様である．

[1]IPv6 ネットワークの運用要件や関連する技術仕様の策定を行うワーキンググループ．`https://datatracker.ietf.org/wg/v6ops/about/`

[2]Stateless IP/ICMP Translation Algorithm

BR

BR[3]とは，SIIT-DC ネットワークにおいて IPv4 インターネットと IPv6 ネットワークとの間で SIIT による IPv4/IPv6 トランスレーションを行う機器[4]である．IPv4 インターネットと IDC 内の IPv6 シングルスタックネットワークの各境界部に所在し，後述する EAMT を参照した 1:1 のアドレス変換を行う．IDC ネットワークに IPv4 インターネットとの接続点が複数ある場合，接続点ごとに最低一つの BR を配備する．

ER

ER[5]とは，IPv4 ネットワークと IPv6 ネットワークとの境界点において多:多の IPv4/IPv6 トランスレーションを行う機器である．SIIT-DC ではそのオプションとして，IPv4 ネットワーク内の IPv4 しか利用出来ないホストが，SIIT-DC を利用して IPv4 サービスを提供するユースケースをサポートする SIIT-DC Dual Translation Mode[23] が定義されており，ER はその中での利用が想定されている．

通常，ER が参照する EAMT 中の EAM は IDC ネットワーク内の IPv4 ネットワーク全体を包括的に指定した IPv4 ネットワークアドレスと，その IPv4 ネットワークを表す IPv6 サービスアドレスにより構成される．

IPv4 サービスアドレス

IPv4 サービスを提供する IPv6 シングルスタックネットワークに属するホストに割り当てる IPv4 アドレス (群) を IPv4 サービスアドレスと呼称する．このアドレス宛に送信されたパケットは，BR/ER によって対応する IPv6 サービスアドレスに変換される．

なお，IPv4 サービスアドレスは BGP[11] によって IPv4 インターネットに経路広告されている必要がある．

IPv6 サービスアドレス

ER/BR を介してアプリケーションやホストに割り当てられた IPv6 アドレス (群) を IPv4 サービスアドレスと呼称する．IPv4 クライアントは SIIT-DC のアーキテクチャをを介して，この IPv6 サービスアドレスが割り当てられたホストと通信することが出来る．

変換プレフィックス

変換プレフィックス[6]とは，RFC6052[19] で定義されたプロトコルに従って全ての IPv4 アドレスをマッピングするために用いられるネットワークプレフィックスが 96bit の IPv6

[3]Border Relay
[4]専用機器もしくは他の役割を有する機器の一機構.
[5]Edge Relay
[6]Translation Prefix.

プレフィックスである．IANA によって主に WKP[7] として 64:ff9b::/48 が予約 [24, 25] されているが，運用者の裁量で ISP 自身に割り当てられた NSP[8] を利用する事ができる．

IPv4 アドレスと IPv6 アドレスの間で変換を実行する際に，B は BR/ER は変換前の IP ヘッダーのアドレスフィールドを、変換プレフィックスが挿入・削除された状態に書き換える。

なお SIIT-DC ネットワークにおいて，変換プレフィックス宛のパケットは各 BR/ER の IPv6 インターフェース宛に IGP[9] などでルーティングされる必要がある．

EAM

EAM[10] とは，EAM アルゴリズム [26] によって結びつけられた IPv4 サービスアドレスと IPv6 サービスアドレスのペア[11] を表す．

EAM において，IPv4 サービスアドレスと IPv6 サービスアドレスは同数[12] である必要がある．

また，BR 及び ER が変換を行う際に参照する EAM 群が記録されたテーブルを EAMT[13] と定義している．以後 EAMT もしくは変換テーブルと呼称する．

3.1.3 ネットワーク設計

基本的な SIIT-DC ネットワークを図 3.1 に示す．IPv6 シングルスタックネットワークにおける IPv4 インターネットとの接続点に BR を置くのみで対外的な IPv4 サービスの提供画家可能になるため，デプロイメントが容易で，様々な IPv6 シングルスタックネットワークにそのまま導入することが出来ることが大きなメリットに挙げられる．

BR は IPv4 インターネットとの各接続点に配置される．各 IPv4 サービスアドレスは BGP により接続先の AS[14] に対して経路広告される．変換プレフィックス宛のパケットは各 BR に広告される．

BR が複数ある場合，それぞれの BR が IGP 等で広告する変換プレフィックスを分けるか，エニーキャスト [27] によって複数の BR が同一の変換プレフィックスを広告するようにする．エニーキャストを使用した場合，BR の障害時に別の BR へとトラフィックを迂回させることが可能になる．

[7] Well Known Prefix.
[8] Network Specific Prefx. 主に RIR から割り当てられた IPv6 Global Unicas Address を指す.
[9] Interior Gateway Protocol
[10] Ecplicit Address Mapping
[11] サービスアドレスはそれぞれネットワークプレフィックスとして指定することも想定されている.
[12] 標準では結び付けられた IPv6 サービスアドレスが IPv4 サービスアドレスより多い状態が想定されているが，IPv6 サービスアドレスのホスト部が若いものから優先して変換するため，余剰分のアドレスは無視される.
[13] Ecplicit Address Mapping
[14] Aunotoumous System. インターネットを構成する自律した組織.

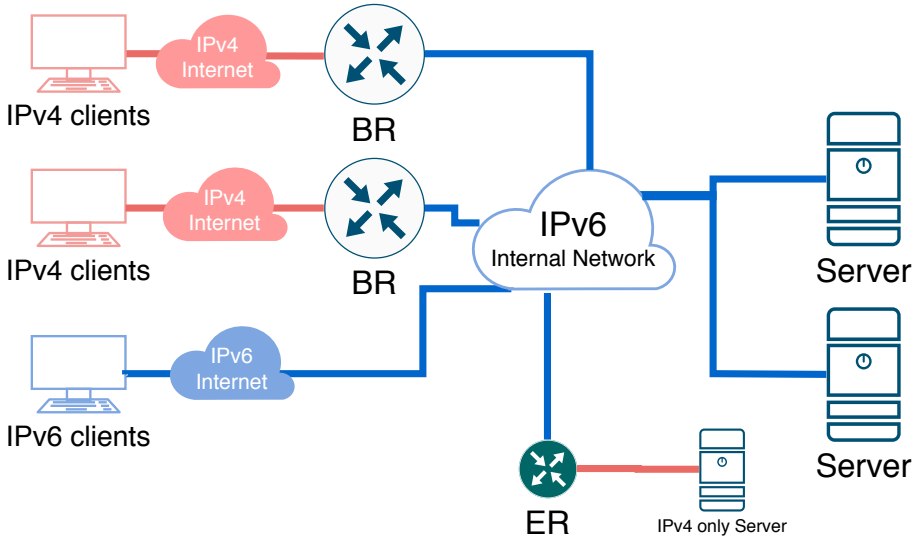

図 3.1: SIIT-DC ネットワーク

ER は IDC 内の IPv4 ネットワークとの接続点に配置され，IPv4 のみを持つホストが IDC 内の IPv6 ネットワークを介して IPv4 インターネットにサービス提供を行う場合に利用される．

3.1.4 基本的なパケットの流れ

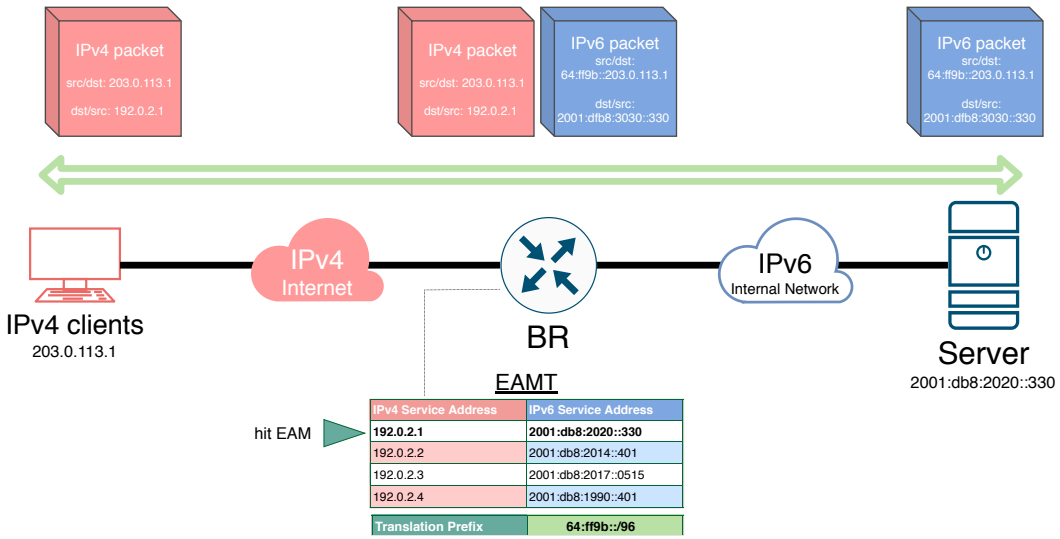

図 3.2: SIIT-DC パケットの流れ

SIIT-DC における基本的な IPv4 クライアントからのトラフィックの流れは以下の様になる．一連のパケットの送信元・送信先のアドレスの遷移を図 3.2 に示す．

IPv4 クライアントの IPv4 サービスアドレス宛のパケットは IPv4 インターネットに接続する BR に到達後，当該 BR が有する EAMT に従って IPv6 サービスアドレス宛の IPv6 パケットに変換される．このパケットの送信元アドレスは変換プレフィックスに埋め込まれた IPv6 アドレスとして表現される．IDC 内の IPv6 ネットワークを介して IPv6 サーバーに到達した後，IPv6 サーバーは送信元アドレスへの応答パケットを送信する．第 3.1.2 項で述べたように，変換プレフィックス宛のパケットは IPv6 ネットワークを経由して BR にルーティングされる．IPv6 サーバーからの応答を受け取った BR は EAMT を参照し，送信元アドレス (IPv6 サービスアドレス) を IPv4 サービスアドレスに書き換え，送信先アドレス (IPv4 クライアントの IPv4 アドレス) から変換プレフィックスを除去書き換えたのち，IPv4 インターネットを介して IPv4 クライアントに返送される．

3.2 SIIT-DC の課題

本節では SIIT-DC の現状の課題及びそれに起因して起こる事象に関して述べる．

3.2.1 一貫した EAMT の必要性

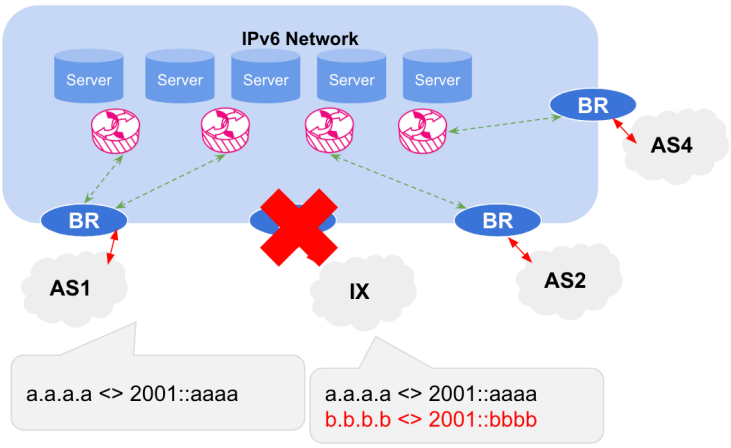

図 3.3: BR に障害が発生した場合に適切にフェイルオーバーが出来ないケース

第 3.1.2 で述べたように，SIIT-DC では対外接続点ごとに BR を配置するネットワークデザインを採用することで，IPv6 シングルスタックネットワークに最小限の IPv4 ネットワークを追加するのみにより IPv4 サービスの提供を可能にしている．また第 3.1.3 項で触れたように，複数の BR で共通した変換プレフィックスをエニーキャストで IDC ネットワーク内に広告する運用を行うことにより，BR 及び対外接続点の障害時に他の BR を用

いて IPv4 サービスの提供を継続することが出来る．この機構を有効に作用させるためには，SIIT-DC ネットワーク内の全ての BR で一貫した EAMT の保持が求められる．

しかしながら現状の SIIT-DC 及び EAMT の仕様 [9, 23, 26] では，BR は他の BR との間で EAMT を共有するためのメッセージング機構を有さない，これは BR 間で EAMT の不一致が発生した場合に，差異となった EAM に該当する IPv4 サービス宛のトラフィックの別の BR への迂回が出来なくなるケースが発生することを意味する．

3.2.2 変更追従性の欠如

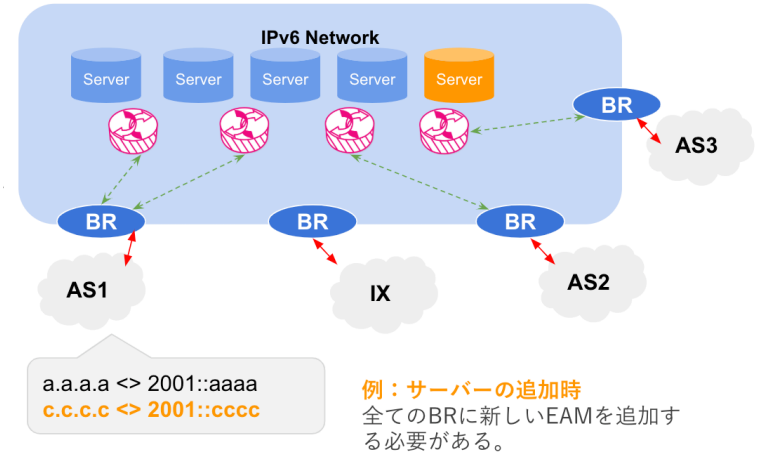

図 3.4: サーバーを追加した際，全ての BR への設定追加が必要になる．

プライベートクラウド環境が一般的に利用される IDC ネットワークでは，日々多くのサーバーやアプリケーションが追加・廃止・変更される．一方で第 3.2.1 で触れたように，SIIT-DC で IPv4 提供サービスを冗長に運用するためには，IPv4 提供サービスに該当する EAM が BR の EAMT に保持されることを要求している．IPv4 提供サービスの構成に変更があった場合，全ての BR の EAMT に対して EAM の更新を行わなくてはならない．

しかしながら現状 SIIT-DC 及び EAMT の仕様 [9, 23, 26] において，IPv4 サービスを行うサーバーの存在や状態によってダイナミックに EAMT を更新する機構は存在しない．そのため，IDC ネットワークにおける IGP などによって IPv6 サービスへの到達性が検証されていたとしても，IPv4 サービスの場合はリアルタイムな構成変更に追従することが出来ない．

第4章　手法の検討

4.1　概要

第 3.2 項で述べたように，現状の SIIT-DC 及び EAMT の仕様は EAMT の一貫性を担保する手法の検討がなされておらず，それに起因した障害時の適切なフェイルオーバーの実行や IPv4 サービスの増減時の変更追従に関しての課題がある．NPO 日本ネットワークセキュリティ協会 (JNSA) らの調査によれば IT システムの障害の原因の約半数は人為ミスに分類されるものにあり [28]，サービスの安定的な稼働を実現するためには単調な繰り返し動作を含む運用をシステムによって減らす必要がある．

本研究では SIIT-DC におけるダイナミック EAMT[1] の実現を目指す．本章では考えられる手法を大別した上でその特徴と利点及び欠点を挙げ，最も適した手法を検討する．

4.2　求められる要件

第 2.1.1 で述べた IPv4 サービス提供手法の機能要件と，第 3.2 項で挙げた SIIT-DC の現状の課題を総合し，EAMT を動的に制御する手法に求められる要件を下記のように定義した．

BR 間の EAMT の一貫性

第 3.2 項で述べたように，障害時の適切なフェイルオーバーを実現するためには，各 BR の EAMT の一貫性が保証される必要がある．

変更追従性

近年の IDC では多数の物理サーバーを統合的に管理するプライベートクラウド環境やコンテナオーケストレーション環境[2] が普及しており，アプリケーション・サービスの追加及び削除が頻繁に行われている．サービスの障害時に適切にそれを検知し，適切に冗長系に移行するはは SLB[3] を中心として広く利用されている．SIIT-DC の IPv4 サービス提供の場でも，サービスの状態の変動にリニアに対応しフェイルオーバーできるような働きが求められる．

[1] BR 群の EAMT をシステムにより動的に制御する機構
[2] Container Orchestration. コンテナ型仮想化統合管理環境
[3] Servver Load Balancer

スケーラビリティ

第 2.1.1 で述べたように IPv6 シングルスタックネットワークにおける IPv4 サービスの提供では水平スケールが容易に行える仕組みを備える必要がある．IPv4 サービスを行うサーバーの増設や，対外接続点が増えた場合の BR の拡大に十分に適用するスケーラビリティを有することが望ましい．

次項ではスケーラビリティの評価のために，制御に必要な通信コネクション数による比較を行う．以後 BR の数を M，IPv4 サービスを提供するサーバーの数を N とし，総通信コネクション数を C として表現する．

4.2.1 デプロイメントの容易さ

第 3.1.3 で述べたように，SIIT-DC の最も特筆すべきメリットの一つにデプロイメントの容易さが挙げられる．これを損なうこと無くダイナミック EAMT が導入されることが望ましい．

4.3 アプローチの分類と比較

本説ではダイナミック EAMT を実現するアプローチとして，二系統のアプローチを考案する．それぞれのアプローチで考えられる実装と実際の構成，及び第 4.2 節で述べた各要件への適合を定性的に評価する．

4.3.1 中央管理型アプローチ

中央管理型アプローチとは，複数の BR の EAMT を統合的に管理する「コントローラー」を IDC ネットワーク上に配置し，各 BR がネットワークを介してこれを参照する機構である．図 4.1 に中央管理型アプローチによってダイナミック EAMT を実現した SIIT-DC の各コンポーネントの関係図を表す．

中央管理型アプローチではコントローラーが各 BR に投入する EAM が記録された「マスターテーブル」を保持し，それを元に各 BR のデータプレーンにルールを書き込む手法を取る．マスターテーブルに記載される EAM はオペレーターがネットワークの構成変更に合わせて追加・削除・更新を行い，それぞれの IPv4 サービスを提供するサーバー群に対してはコントローラからプル型[4]の外部監視[5]によりサーバーの状態変化を検知しマスターテーブルを更新する．

本アプローチの実装手法としては，OpenFlow[6]を用いた集中コントローラー型 SDN フレームワークを利用する方法が考えられる [29]．類似事例として，Sheng らによって Open

[4]pull-based monitoring. コントローラから各サーバーに能動的に情報を取得する

[5]External monitoring

[6]Open Networking Foundation により標準化されているデータプレーン制御用通信プロトコル．`https://www.opennetworking.org/`

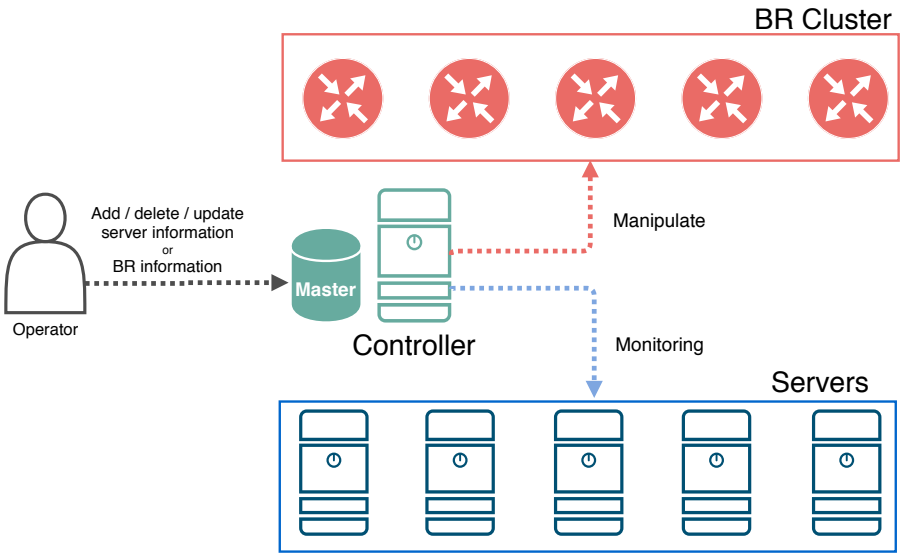

図 4.1: 中央管理型アプローチによるダイナミック EAMT

Flow を利用して各アクセススイッチに IPv4/IPv6 トランスレーション機構をデータプレーンとして導入するデータセンターネットワークデザインの提案がなされている [30].

要件評価

- BR 間の EAMT の一貫性
  本アプローチでは各 BR の EAMT が一つのマスターテーブルからレプリケーションされるために，十分な一貫性が保証される.

- 変更追従性
  基本的には EAM 情報の更新はオペレーターのマスターテーブルへの記入までの時間はコントローラーのサーバー監視性能に依存する.

- スケーラビリティ
  コントローラーの数を L とすると，EAMT の制御に必要とする総通信コネクション数 C は以下の通りになる.

$$C = L(M + N) \tag{4.1}$$

  一方，変更追従性と同じくどこまでの筐体を収容できるかはコントローラーの実装・性能がボトルネックになる設計となる.

- デプロイメントの容易さ
  コントローラーに求められる機器の性能・機能要件が大きいため，標準的な SIIT-DC よりデプロイメントのコストは向上する.

4.3.2 分散管理型アプローチ

分散管理型アプローチとは，IPv4 サービスを提供するサーバーがエージェントプロセスを介して自身の IPv4 サービスアドレスと IPv6 サービスアドレスを広告し，その広告情報を受け取った BR が自身の EAMT に反映させる機構である．図 4.2 に中央管理型アプローチによってダイナミック EAMT を実現した SIIT-DC の各コンポーネントの関係図を表す．

サーバー群は各 BR と EAM を広告するためのコネクションを確立する．IPv4 サービスを提供するサーバーと BR の間の IP ネットワークが何らかの原因により疎通不能になると，当該サーバーの広告も同時に停止されるため，該当 BR の EAMT から該当する EAM のレコードが削除される．

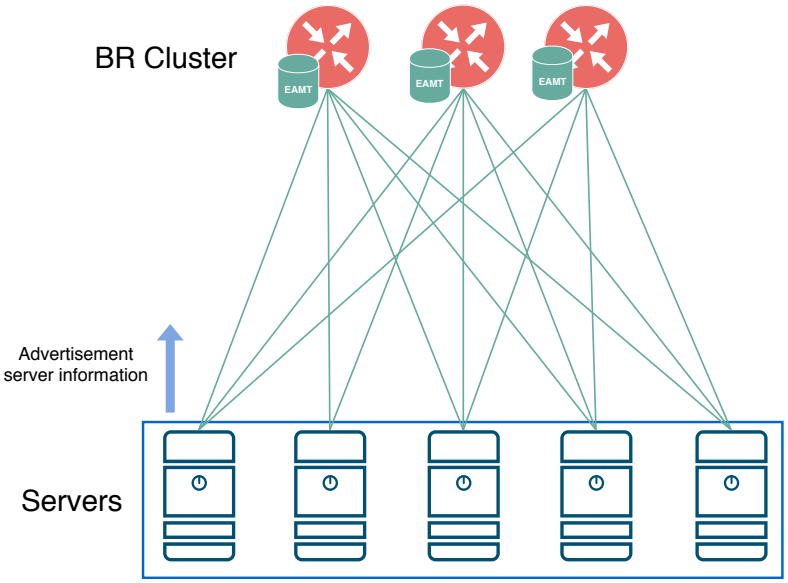

図 4.2: 分散管理型アプローチによるダイナミック EAMT

要件評価

- BR 間の EAMT の一貫性
  各 BR 間で EAMT 一貫性を保証する機構は無いが，当該 BR と疎通できないサーバーは障害時に自身の IPv4 サービスアドレう宛のトラフィックを当該 BR に経由させることが出来ないため，問題にならない．

- 変更追従性
  サーバー自身のエージェントプロセスが直接 BR に広告を行うため，実際の変更にリニアに対応出来る．

- スケーラビリティ
  EAMT の制御に必要とする通信コネクション数 C は以下の通りになる．

$$C = M \cdot N \tag{4.2}$$

  サーバー群・各 BR 間でフルメッシュでのコネクションが必要なため，SIIT-DC ネットワーク自体が小規模の場合のみ採用可能である．

- デプロイメントの容易さ
  各サーバー・BR にエージェントを導入する必要があるが，システム自体の機能は軽量である．

4.4 アプローチの検討

中央管理型アプローチが各 BR 間での EAMT 一貫性，スケーラビリティの二要素で優位であるが，コントローラーの役割が非常に大きくなり機能要件が高くなるため，変更追従性とデプロイメントの容易さの面での障壁が高いという問題も抱えている．一方で分散管理型アプローチはシンプルな構成であるためデプロイメントが比較的容易であり変更への追従がリニアであるが，各サーバーが通信コネクションを多量に貼らなくてはならない点でスケーラビリティに難がある．

本提案手法では両アプローチを総合した動的経路制御プロトコルである iBGP を利用したハイブリッド型アプローチを提案する．第 5 章において，本提案手法を中央管理型・分散型の両アプローチと比較する．

第5章　提案手法

本章では提案手法の設計を述べる.

5.1　概要

第6章　プロトコル設計と実装

本章では，提案システムのメッセージ設計と実装について述べる．

6.1　実装内容

第7章　評価

本章では，本研究の評価を行う

7.1　評価要件

第8章　結論

本章では，本研究のまとめと今後の課題を示す.

8.1　本研究のまとめ

8.2　本研究の課題

謝辞

俺に関わった全てに感謝

参考文献

[1] potaroo. Ipv4 address report. https://ipv4.potaroo.net/. 最終閲覧: 2019-12-17.

[2] Cisco. Cisco visual networking index: Forecast and trends, 2017–2022 white paper, 2017. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html.

[3] Felipe Alonso and John Boucher. Business continuity plans for disaster response. *The CPA Journal*, 71(11):60, 2001.

[4] 石田慶樹, 吉田友哉, and 西田圭. 日本のインターネットは本当にロバストになったのか？ In *JANOG 44* ミーティング, 2019. https://www.janog.gr.jp/meeting/janog44/application/files/7715/6577/5523/janog44-robust-ishida-01.pdf.

[5] IANA. Internet protocol version 4 address space. https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml, 2019. 最終閲覧: 2019-12-17.

[6] Lee Howard and Time Warner Cable. Internet access pricing in a post-ipv4 runout world. *White Paper*, 2013.

[7] Alain Durand. Deploying ipv6. *IEEE Internet Computing*, 5(1):79–81, 2001.

[8] Google. Ipv6 statistics. https://www.google.com/intl/en/ipv6/statistics.html. 最終閲覧: 2019-12-18.

[9] Tore Anderson. SIIT-DC: Stateless IP/ICMP Translation for IPv6 Data Center Environments. RFC 7755, February 2016.

[10] Congxiao Bao, Xing Li, Fred Baker, Tore Anderson, and Fernando Gont. IP/ICMP Translation Algorithm. RFC 7915, June 2016.

[11] Yakov Rekhter, Susan Hares, and Tony Li. A Border Gateway Protocol 4 (BGP-4). RFC 4271, January 2006.

[12] Jordi Palet, Hans M.-H. Liu, and Masanobu Kawashima. Requirements for IPv6 Customer Edge Routers to Support IPv4-as-a-Service. RFC 8585, May 2019.

[13] A. Vahdat, M. Al-Fares, N. Farrington, R. N. Mysore, G. Porter, and S. Radhakrishnan. Scale-out networking in the data center. *IEEE Micro*, 30(4):29–41, July 2010.

[14] Katja Gilly, Carlos Juiz, and Ramon Puigjaner. An up-to-date survey in web load balancing. *World Wide Web*, 14(2):105–131, Mar 2011.

[15] Patrick Shuff. Building a billion user load balancer. Dublin, May 2015. USENIX Association.

[16] Daniel E. Eisenbud, Cheng Yi, Carlo Contavalli, Cody Smith, Roman Kononov, Eric Mann-Hielscher, Ardas Cilingiroglu, Bin Cheyney, Wentao Shang, and Jinnah Dylan Hosein. Maglev: A fast and reliable software network load balancer. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 523–535, Santa Clara, CA, 2016.

[17] N. Chuangchunsong, S. Kamolphiwong, T. Kamolphiwong, R. Elz, and P. Pongpaibool. Performance evaluation of ipv4/ipv6 transition mechanisms: Ipv4-in-ipv6 tunneling techniques. In *The International Conference on Information Networking 2014 (ICOIN2014)*, pages 238–243, Feb 2014.

[18] Philip Matthews, Iljitsch van Beijnum, and Marcelo Bagnulo. Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers. RFC 6146, April 2011.

[19] Xing Li, Mohamed Boucadair, Christian Huitema, Marcelo Bagnulo, and Congxiao Bao. IPv6 Addressing of IPv4/IPv6 Translators. RFC 6052, October 2010.

[20] Christian Hopps. Analysis of an Equal-Cost Multi-Path Algorithm. RFC 2992, November 2000.

[21] Erik Nordmark. Stateless IP/ICMP Translation Algorithm (SIIT). RFC 2765, February 2000.

[22] Xing Li, Fred Baker, and Congxiao Bao. IP/ICMP Translation Algorithm. RFC 6145, April 2011.

[23] Tore Anderson and S.J.M. Steffann. Stateless IP/ICMP Translation for IPv6 Internet Data Center Environments (SIIT-DC): Dual Translation Mode. RFC 7756, February 2016.

[24] Tore Anderson. Local-Use IPv4/IPv6 Translation Prefix. RFC 8215, August 2017.

[25] IANA. Internet protocol version 6 address space. `https://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml`, 2019. 最終閲覧: 2019-12-17.

[26] Tore Anderson and Alberto Leiva. Explicit Address Mappings for Stateless IP/ICMP Translation. RFC 7757, February 2016.

[27] Kurt Erik Lindqvist and Joe Abley. Operation of Anycast Services. RFC 4786, December 2006.

[28] NPO 日本ネットワークセキュリティ協会 (JNSA). 情報セキュリティインシデントに関する調査報告書. https://www.jnsa.org/result/incident/2018.html, 2018. 最終閲覧: 2019-12-21.

[29] Evangelos Haleplidis, Kostas Pentikousis, Spyros Denazis, Jamal Hadi Salim, David Meyer, and Odysseas Koufopavlou. Software-Defined Networking (SDN): Layers and Architecture Terminology. RFC 7426, January 2015.

[30] S. Maojia, B. Congxiao, and L. Xing. A sdn for multi-tenant data center based on ipv6 transition method. In *2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference*, pages 190–195, May 2016.