第14部

ネットワーク相互接続の実証実験
Programmable Internet Exchange in EDO (PIX-IE)

関谷 勇司

第１章　はじめに

本研究では、商用インターネットを相互に接続する場合の問題点を明確にし、それを解決するための技術や手法の研究開発ならびに実証実験を行う。特に、近年成長し続ける動画系のインターネットトラフィックや、スマートフォンのファームウェア更新などによる突発的なトラフィック増大、スマートフォンアプリの流行にともなう一時的なトラフィック増大等の傾向に対して、トラフィックの輻輳を防ぎ、ユーザへの応答性を保つためのトラフィックエンジニアリング手法の検討と検証を行う。また、大規模災害等の障害にも対応できるための強固なインターネットバックボーンの形成に関する実証実験を行う。

さらに、特に注力している研究テーマとして、Software Defined Network (SDN) 技術のIXへの導入があげられる。SDN技術をIXに導入することにより、トラフィックの柔軟な制御や攻撃を防御するためのセキュリティ機能を提供することができる。本報告書では、特にこのSDN技術を用いた次世代IXである、PIX-IE(Programmable Internet Exchange)の実現を目指した研究活動について述べる。

本研究は、WIDE Projectのサブプロジェクトである Network Service Provider Internet exchange Point (NSPIXP) プロジェクトとして行われている。NSPIXPプロジェクトは、日本初のIXを構築・運用したプロジェクトであり、現在はDIX-IE、NSPIXP-3、NSPIXP-23と呼ばれるIXを運用し、インターネットがより信頼性を有した高度情報インフラストラクチャとして機能するために必要となる機能の検証や開発、ならびにその実証実験を行っている。PIX-IEはこれらのIXに続く、実験的なIXとして構築・運用されている。

本報告書では、第2節にてプロジェクトの背景と現在の構成を述べ、第3節にて本年度の研究成果を報告する。最後に第4節にてまとめとこれからの展望について述べる。

第２章　プロジェクトの背景と現状

NSPIXPプロジェクトは、1994年のNSPIXP-1運用開始、1996年のNSPIXP-2運用開始、1997年のNSPIXP-3運用開始を経て、現在は、東京エリアに分散配置されたDIX-IEと、大阪に配置されたNSPIXP-3、ならびにこの2つのIXを結合した、NSPIXP-23、SDN技術を導入したIXであるPIX-IEという、4つのIXを運用している。全てのIXはIPv4/IPv6デュアルスタックにて運用されている。表1に平成30年1月時点での、各IXの実証実験拠点を示す。

表1　各IX拠点一覧

DIX-IE	KDDI 大手町拠点
	NTT コミュニケーションズ大手町拠点
	ComSpace-1 拠点
	@Tokyo 拠点
	NTT Data 大手町拠点
NSPIXP-3	NTT テレパーク堂島拠点
NSPIXP-23	KDDI 大手町拠点
	NTT コミュニケーションズ大手町拠点
	NTT テレパーク堂島拠点
	ComSpace-1 拠点
	@Tokyo 拠点
	NTT Data 大手町拠点
PIX-IE	KDDI 大手町拠点
	NTT コミュニケーションズ大手町拠点
	NTT Data 大手町拠点
	NTT テレパーク堂島拠点

また、平成30年1月時点での、DIX-IEならびにNSPIXP-23の構成トポロジを図1に示す。同様に、平成30年1月時点における、PIX-IEの構成図を図2に示す。PIX-IEは実験的IXであるため、構成図に利用しているスイッチの機種名も明記した。紫色の枠にて囲まれているスイッチが、PIX-IEを構成しているOpenFlowスイッチとなる。

PIX-IEにおけるOpenFlowコントローラは、東京エリアと大阪エリアでそれぞれ設置している。これは、コントローラ間の通信に障害が発生した場合、遠隔拠点のOpenFlowスイッチが制御できなくなることを防ぐため、それぞれのエリア(東京・大阪)単位でOpenFlowコントローラを設置する構成を選択した。

このように、拠点障害に対応するための分散IXアーキテクチャの構築と、高信頼性を実現するための冗長化IXアーキテクチャの構築と運用に関する実証実験を行っている。

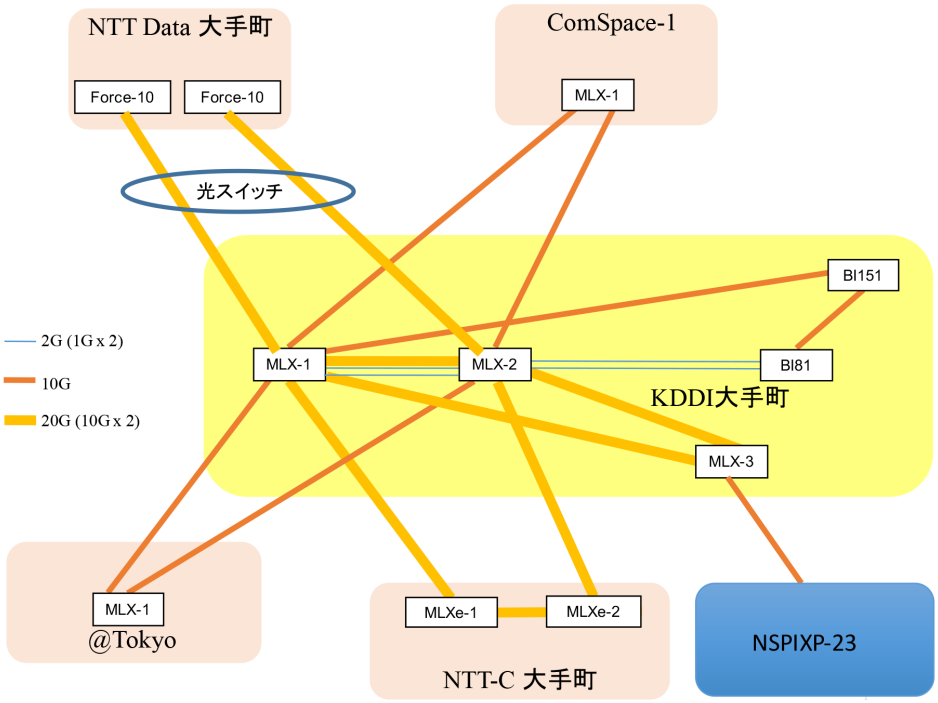

図1　DIX-IEならびにNSPIXP-23構成図

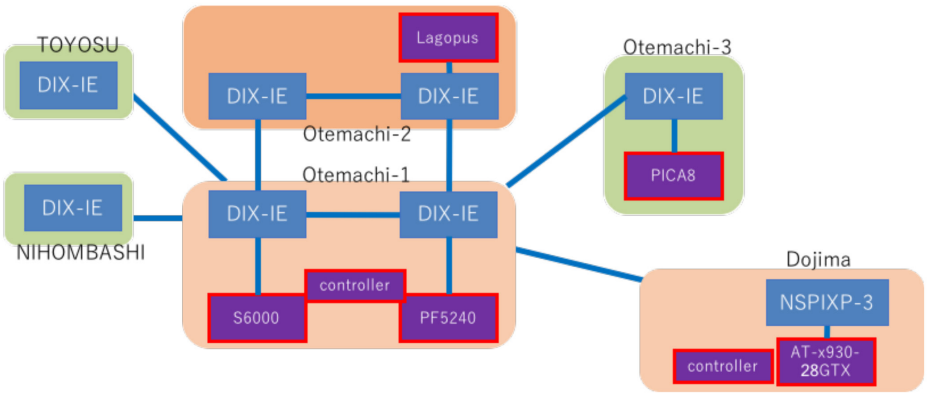

図2　PIX-IE構成図

第３章　研究成果

3.1　新たなトラフィック傾向に対応したIXサービスアーキテクチャの研究

近年のスマートフォンの普及や動画コンテンツによるトラフィックの増大、ならびにIoTに代表されるような新たなトラフィックの動向に対して、コストバランスを持って対応することのできるIXの構築に関して取り組んだ。ポイントとしては、従来のIXのような高価かつ大規模な装置を使うのではなく、低価格かつ小規模な装置を連結することで、広域IXを構成することを目指した。さらに、前述のような多種のトラフィック動向に対して、BGPによる経路制御のみならず、より細かな粒度でトラフィック制御を行うためにOpenFlow技術を導入したIXの構築に取り組んだ。

その結果として、前年度の報告書においても報告した通り、東京エリアの複数拠点においてPIX-IEと呼ばれるSDN技術を用いたIXを構築し、運用を開始した。このPIX-IEにて利用している機材は、以下の通りである。

- DELL S6000-ON (KDDI大手町拠点)
- NEC PF5240 (KDDI大手町拠点)
- Lagopusソフトウェアスイッチ (NTTコミュニケーションズ大手町拠点)
- Pica8 (NTT Data大手町拠点)
- Allied Telesis AT-x930-28GTX (NTTテレパーク堂島拠点)

どの機材も1Uサイズのスイッチ、もしくはサーバ機材を利用したソフトウェアスイッチであり、従来IXにて利用されていた機器より小型で低価格なものとなっている。これらを連結して論理的に1台のL2スイッチを構成し、かつ必要なトラフィックのみが疎通するよう構築されたものがPIX-IEである。

また、NTTテレパーク堂島拠点は、新たに平成29年7月にPIX-IE拠点に追加された。この拠点にて利用されたスイッチはAllied Telesis社のスイッチであり、初の導入となる。NTTテレパーク堂島拠点において設置されたPIX-IEスイッチの様子を図3に示す。図中の赤丸にて囲まれた1Uサイズのスイッチが PIX-IEのスイッチであり、その上部にマウントされているBrocadeのスイッチがNSPIXP-3となっている。

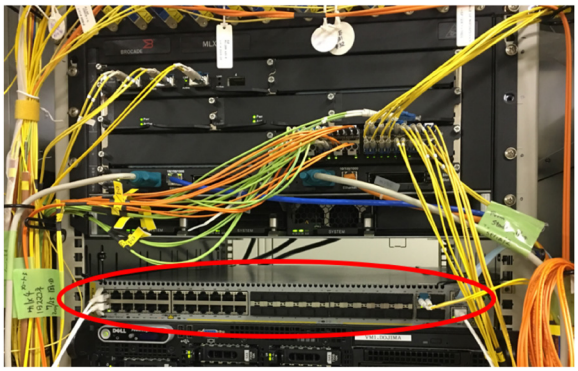

図3　PIX-IE NTTテレパーク堂島拠点

さらに、PIX-IE NTTテレパーク堂島拠点では、制御のためのSDNコントローラに新たなソフトウェアを導入した。東京エリアでのPIX-IEでは、NSPIXP Projectによって自作されたコントローラを利用していた。しかし、自作のコントローラの場合、利用するOpenFlowスイッチの種類が増えた場合に自身で動作を検証する必要があり、新たな機能を実装しようとした場合にもその機能を一から作成する必要があった。そのため、今回新規導入したPIX-IE NTTテレパーク堂島拠点では、FAUCET[1]と呼ばれるオープンソースのSDNコントローラを導入した。FAUCETはPython言語にて記述されており、OpenFlowスイッチにてファブリックを構築するためのSDNコントローラである。複数種類のOpenFlowスイッチをサポートし、それらを連結して論理的に1台のOpenFlowスイッチやL2スイッチ、L3スイッチを構成することができる。FAUCETの公式Webページによると、以下の機能がサポートされている。

- VLANs
- IPv4 and IPv6 support
- IPv6 neighbor solicitation and router advertisement support

＊1　http://faucet.nz/

- Static and BGP routing
- Flexible port and VLAN based Access Control Lists
- Port mirroring
- Fast configuration reloads
- Vendor neutral stacking of Openflow switches
- Policy based forwarding to offload processing to external systems (Eg 802.1x via hostapd)
- Configurable learning: Control unicast flooding by port and by VLAN
- Dataplane for NFV - Offload functions such as DHCP, NTP, Firewall, and IDS
- CouchDB support for storing flows from switches to enable north bound applications
- Influx support for time-series OpenFlow port statistics
- Prometheus integration for monitoring and instrumentation of FAUCET
- Grafana based dashboards for monitoring

すなわち、OpenFlowスイッチを利用して、通常のL2スイッチやL3スイッチが有する機能を手軽に実現することができる。この中でIXの構築に利用できる機能は少ないが、このFAUCETの上にTouSIX Projectにて開発されたUmbrella[2]という方式を実装することで、OpenFlowスイッチを用いたIXを実現した。Umbrellaに関しての詳細は論文[3] “ENDEAVOUR : A Scalable SDN Architecture for Real-World IXPs” に述べられている。今回この方式の作者である、Marc Bruyere研究員をNSPIXP Projectに迎え入れ、PIX-IEにて利用する新たなSDNコントローラを構築した。

Umbrella方式の利点は、各OpenFlowスイッチやSDNコントローラが通信の状態を保持することなくIXに接続されたユーザ同士のトラフィック交換が行えることである。現在東京エリアで用いられている独自開発のPIX-IE SDNコントローラにおいても、各スイッチは基本的に静的なルールのみを利用しているが、ARPやNDPの処理にOpenFlowのPacketIN、PacketOUTという処理を用いているため、OpenFlowスイッチとSDNコントローラの間での通信が頻繁に行われる。その点Umbrella方式では、ARPやNDPも静的ルールのみで解決されるため、不必要なARPやNDP、事故によって発生する突発的なブロードキャストやマルチキャストを除外しながら、必要なARP, NDPトラフィックを確実に伝達することができる。すなわち、IXでの通信事故を防いだ、より信頼性の高いIXを構築することができる。また、FAUCETの上にUmbrella方式を実装することにより、OpenFlowスイッチの種別を意識することなく、FAUCETの機能を利用してより簡易にPIX-IEの付加機能を実現することが可能となるためである。また、このFAUCETをPIX-IE堂島拠点に導入するにあたって、以下のテスト項目を実施した。

- ARP unicast試験
- ICMPv6 ND unicast試験
- ルール不適合パケット破棄試験
- IPv4/IPv6 unicastトラフィック試験
- ARP/NDを送信しながらのスループット試験
- FAUCETコントローラ再起動時におけるトラフィック挙動試験

これらのテスト項目をすべてパスしたOpenFlowスイッチは、LagopusとAlliedTelesis x930であった。そのため、PIX-IE堂島拠点にAllied Telesis x930を導入した。

NTTテレパーク堂島拠点に構築されたFAUCET SDNコントローラの構成図を図4に示す。GAGUGEはOpenFlowスイッチからトラフィック情報などの統計情報を収集するモジュールであり、Grafanaは可視化ツールである。

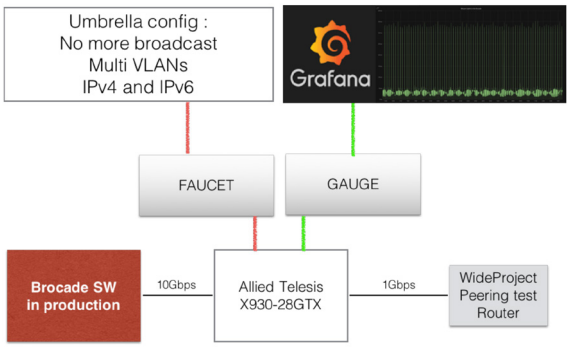

図4 FAUCET SDNコントローラ構成図

＊2 https://blog.apnic.net/2017/05/08/tousix-project-sdn-ixps-design-production/
＊3 http://www.dia.uniroma3.it/~compunet/www/docs/chiesa/endeavour-jsac.pdf

現在、NTTテレパーク堂島拠点のみこのFAUCET + UmbrellaによるSDNコントローラを利用してPIX-IEが運用されているが、本年度後半から来年度にかけて、東京地区のSDNコントローラもFAUCET + Umbrellaに置き換えることを計画している。

3.2 IXにおけるセキュリティに関する研究

現在のインターネットにおいて、サービス妨害攻撃は管理者や運用者を悩ませる大きな課題の一つである。前年度の報告書においては、DDoSの発生事例とIXにおけるその防御手法に関しての検討結果を述べた。今年度の研究期間中には、ユーザの経路リークにより一部サイトへの到達性が不安定となる事故が発生した。この事故を事例として、IXにおいてどのような機能を提供すればより信頼性の高いIXが構築できるかを検討した。

平成29年8月25日にこの経路事故は発生した。AS15169が通常では広告していない、他ASが起源となる経路を広告したことにより発生したことが観測されている。この間に観測された、DIX-IEにおけるBGP Updateの状況を図5に示す。

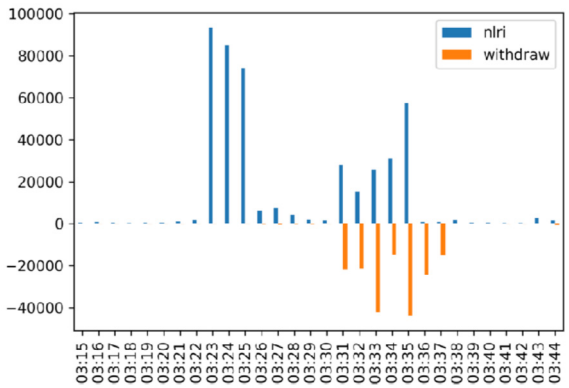

図5 DIX-IEにおけるBGP Update頻度(UTC)
(IRS27のコーダンス小島氏の資料より引用)
http://irs.ietf.to/wiki.cgi?page=IRS27

IRS27(Inter-Domain Rouring Security)の会合にて発表された小島氏の資料によれば、この時約9万経路の増加が観測されたとのことで、この9万経路に該当した一部サイトへの到達性が不安定になったと考えられる。この経路障害自体は、今回は10分程度で収束したと観測されるが、9万経路が瞬時に増加したことにより、ルータに障害をきたした接続ユーザも存在する。そのため、実際の障害は10分間のみならず、その後も影響が続いたものと考えられる。

このような事故を未然に防ぐためには、IX自体にどのような機能があればよいか考察してみた。IXに接続している事業者は、なるべくBGPの運用コストを下げたい傾向にあり、BGPピア単位で細かいフィルタを設定することは難しい。そこで、IXにて自動的な経路フィルタリングを提供する手法を考えた。

- IXにおける広告経路ポリシーの登録
- 一定時間内における急速な経路増加の検知と防御
- BGPのみに頼らないトラフィック交換

まず、広告経路ポリシーの登録は、IX接続者がPrefix単位のフィルタ、もしくはAS-Path単位のフィルタにより、広告すべき経路のポリシーを登録することで、IX側にて伝達する経路を制御する方式である。これを実現するためには、SDNの機能を用いてBGPのピアリングセッションをIXにて中継するか、BGPメッセージの監視を行う必要がある。ポリシーに従わない経路をフィルタリングする場合には、BGPセッションをIXにて中継する必要がある。しかし、IX側のSDNコントローラですべてのBGPセッションを中継する必要があり、SDNコントローラの負荷が高まる。また、接続者によってはポリシーを明記してくれない場合もある。この手法はBGP Routeサーバを用いて経路フィルタリングを行う手法とほぼ同等であるが、組織同士の直接BGPピアリングにも対応できる点が異なる。

次に、一定時間内における急速な経路増加の検知であるが、これもBGP Routeサーバを用いても実現可能である。しかし、急速な経路増加を検知した場合にもBGPセッションを強制的に切断することは難しい。そのため、現実的に効果のある手法とは言えない。

最後に、BGPのみに頼らないトラフィック交換は、SDNの機能を有効に利用できる手法であると考える。BGPではなくSDNコントローラへのAPI発行にて経路登録や受信を行う手法である。この場合、経路単位で認証を行う

といった細かな制御が可能となる一方、既存の経路制御プロトコルとの互換性を確保しなければならないため、IXの接続者側で対応する必要がある。もしくはBGPへの変換を含めてIX側にて提供する手法が考えられる。BGPにて広告した経路がSDNコントローラにて認証されることで、始めて相手組織のBGPに伝達される、という手法である。同様な手法として、RPKIやBGPsec等の技術が存在する。これは利用者側での対応が必要であり、普及という点ではIX内部のみで完結する手法の方が導入しやすいと考えられる。

第４章　まとめ

本報告書では、平成29年度におけるNSPIXPプロジェクトでの研究開発と実証実験に関して、その成果をまとめた。特にPIX-IEに関して、次世代IXを目指す注力技術として、新たなコントローラ導入を含めた研究開発と実証実験を行った。NSPIXPプロジェクトでは、これからのISPやコンテンツ事業者に求められる、高度情報インフラストラクチャとしてのIXサービスのありかたを常に念頭におき、より強固なインターネットバックボーンとサービスを実現するための、高度な運用技術の研究開発ならびに実証実験を行っていく所存である。

特に、PIX-IEは引き続き本研究における最重要テーマであり、その実現に関して最優先に取り組んでいく所存である。安定性と機能性、そして安価なコストを実現した次世代IXを、世界規模での運用に発展させることが、NSPIXPプロジェクトの社会貢献であり、存在意義であると考える。