Internet Week 2025
不確実性に挑むこれからのピアリング運用のキーポイント

IIJ
Internet Initiative Japan

26th Oct 2025

株式会社インターネットイニシアティブ
蓬田 裕一 (よもぎた ゆういち)
y-yomogita@iij.ad.jp

経歴

- 蓬田 裕一 (よもぎた ゆういち)
- 2008-2014 IIJ
  - トランジットサービス運営
  - バックボーン運用
- 2015-2019 JPNAP
  - IXサービスの運営・設計・構築・運用
- 2019- IIJ
  - IIJバックボーンチームマネージャー(AS2497)
  - Peering/Interconnection マネージャー

今日のおはなし

- AS運用のリアルな現場からAS間のピアリング/相互接続のキーポイントをお話します
- 相互接続インタフェースのトレンドやIXP/PNIの考え方
- Peerの運用やDDoS対策の対策例
- さらに今後の相互接続でのキーファンクションとなる機能の紹介を行います

バックボーン

- バックボーンネットワーク
  - 新規POP設置・移設・廃止
  - 新機種選定、検証、新機能導入
  - インターネット接続系サービス設備の運営
  - 内製アプリ開発のProject Manager
- 日々の構築・運用業務
  - AS2497の機器オペレーションや障害対応
  - 運用フローの設計や導入
  - 承認業務

対外対応

- Peering戦略検討と交渉
  - 顧客、ニーズを意識した戦略を検討
  - Peer接続の交渉
- 社外交渉
  - Peering以外にも、海外回線調達、海外機器の保守事業者選定、データセンター選定
- 社外コラボレーション
  - 他社とのサービス協業の検討

Peer/相互接続状況のトレンドと考え方

- **10Gインタフェースの利用減少、移行の傾向が増加**
  - コスト効率の悪化とポート利用効率低下が背景
  - Link Aggregationで束ねていた10G接続は機器更改に合わせて100Gへ移行
- **100GインタフェースがPNIの主力に**
  - 100Gポート搭載ルータの敷居が下がってきた印象
    - 逆にAS borderで利用されるルータでの10G/1G利用の敷居が高くなりつつある
  - 事業者間で100G化の合意形成が進展
  - IXPでの10G→100G移行が加速
  - 100Gの種別は引き続き100G-LR4が中心
  - Single Lambda 100G-LRは欧米IXPを中心にReady、利用は限定的か
    - 移行のモチベーションはコスト低下・部品点数減による信頼性向上
- **400Gインタフェースは限定的利用**
  - 大規模事業者やIXPで検討進むが、一般事業者ではまだ導入途上
  - 400G-FR4 / 400G-LR4 での接続

- **一般的な図式**

要素	Private Network Interconnect	Internet eXchange Point
物理ポートの数	多くなりがち	集約効果あり
Peeringの交渉の難易度	求められる要素が多く、難易度が上がりがち	なんとなく敷居が低い認識
接続のしやすさ	1本1本ハンドメイドでの作成	一度物理接続を作れば、あとは論理設定を入れるだけ
導入までの手間	意外に大変な要因のすり合わせ	IXPにより手厚くサポート
リードタイム	接続回線の納期に引っ張られがち	導入プロセスを経る必要あり
コスト	対ASへの最適化を考えれば最安 でも本数が増えるとコストは指数関数的にあがる	一般的に接続ポートのコストが高く見えがち
運用	影響範囲や制御も1つのASに対してなのでシンプル	Peerの数が多くなったときにインパクトが大きい
品質	向上するでしょう。放置されることも少ない	IXPの品質に依存するところあり
セキュリティ	2者間の取り決め次第。第3者が介入しない	Multilateralによる一抹の不安定さはある

- **私が考えるIXPとPNIの利用方法**
  - IXPの集約効果、接続のしやすさとリードタイムは魅力的
    - 国外の細かいトラフィックはIXPで集約する
    - 複数のDCにまたがるIXPを活用できるのは魅力的
  - トラフィックが変動する、量が多い場合はIXPからPNIへ分割
    - 最近だと100G複数本のLAGを作るくらいなら、PNIにするくらいの感覚

- **Peeringの根底**
  - そこに契約書が存在しない世界 (昔は覚書あった？)
  - 相手と自分がwin-winの関係でなければ意義が存在しない
  - あくまで対等な関係での利益の創出になる
    - 一方的な利益供与は関係がうまくいかなくなる

- **一部の人達はPeeringの意義を改めて考え始めている**
  - なぜPeeringを行うのか、何を目的にするのか
  - 自分たちのサービス戦略にPeeringが適合するのか
  - Peeringに積極的な組織とPeeringを整理し始めている組織の存在
  - トラディショナルなPeeringは整理されていく可能性あり

- **不特定多数とPeeringすることよりも信頼する事業者と接続**
  - インターネットがより社会インフラとしての信頼度を求められる
  - Peeringの一番の目的は効率的、的確なトラフィック交換
  - IIJはポリシーとしてルートサーバは使っていない (接続先を見極めている)
  - Peeringはトラフィックを交換するためではなく、事業を発展させるツール

- **Peerの監視してますか？**
  - 数が多くなると切れる頻度も多い
  - メンテナンス、障害対応を律儀にやる組織もあればやらない組織もある
  - 気づかずしばらくBGPピアが落ちている接続も。

- **あくまで対応の一例**
  - 優先度は決める (Rank S, A, B, C)
    - 自分たちのビジネス/サービスへのインパクトを加味することは重要
  - 監視項目も対応も
    - BGP Peer down, 利用帯域, 経路数, インタフェースカウンタ
    - 検知後 即時対応 / 30分様子見して問い合わせ / 静観

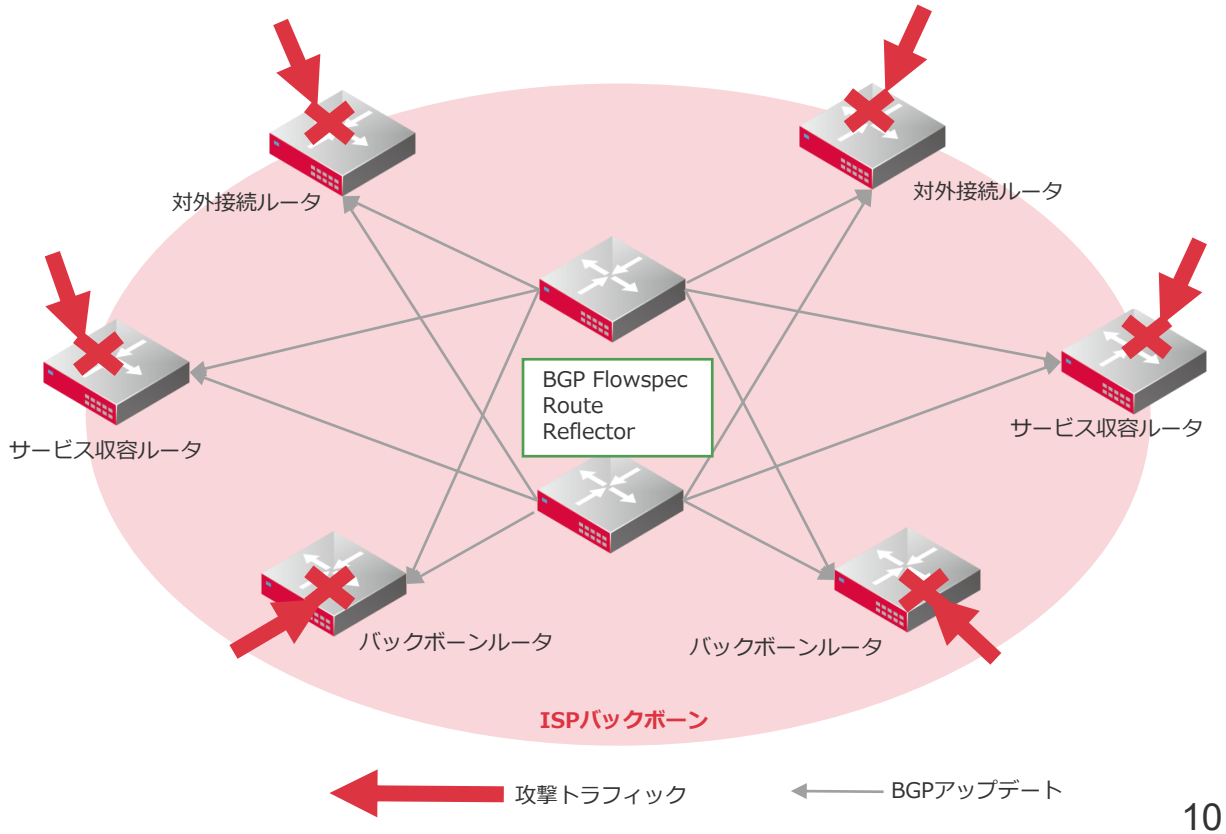
- **経路フィルタ**
  - 一時期盛んだったAS-Path調整業務も今はAS-Pathをかけることも少なくなってきた印象
    - Peerを信頼
    - RPKI ROVによるOrigin Validation
    - IRR as-setから自動生成をやっているところもありそう
  - 怖いのはfull routeの受信および経路リーク。IIJでの防御策はmax-prefixによる制御
    - しきい値を超えるものは即時BGPピアを切断
    - 経路の正当性はしきい値を見直すことで対応

DDoSへ対応するために

- **通信事業者として、顧客通信を侵害しないのが大前提**
  - サービスへの影響がある場合に限り、必要最小限の対処が許される
    - 電気通信事業者におけるサイバー攻撃等への対処と通信の秘密に関するガイドライン
    - https://www.jaipa.or.jp/other/mtcs/guideline_v6.pdf
- **不正なトラフィックかどうかは当事者しかわからない**
  - あくまでDDoSっぽいというだけ
  - どう見てもDDoSだとしても、停めずに運び切るのが大原則
- **まずは、経路制御により余裕のある回線に迂回する**
  - できる限り通信制御をせずに最適な経路から変えることになり遅延に繋がるため迂回は局所的できることを考える
- **とはいえ設備にも限界はあるので、最終手段として通信制御**
  - 発生している対象通信をピンポイントで対処するのが原則
  - ここのところ多いカーペットボミングで/24を舐め回すように攻撃していくなど影響拡大が予期できる場合は対処に苦慮

- **BGP Flowspec(RFC5575)を活用したバックボーン全体での通信制御**
  1. Route ReflectorでFlowspec経路(= 制御ルール)を生成
  2. iBGPアップデートにより各ルータへ制御ルール配布
  3. 各ルータで制御ルールに従った制御の実施

- **Flowspecによる細かな通信制御**
  - source/destination prefix
  - protocol
  - source/destination port
  - action
    - discard、rate-limit
  - communityによる発動先の制御
    - 地域やルータを指定可能
  - 旧来のBlackholeより自由な制御が可能
    - Blackholeは宛先IPのみを指定し全破棄



- **DDoSを受けると相互接続の接続が輻輳することが辛い**
  - 各AS内に入ってきたものは、いかようにでも制御できると思うが、相互接続点のDDoS対策は結構悩ましい

- **対策する手段**
  1. とにかく相互接続点を増強する
  2. 自ASに入ってくる前にトラフィックを何とかする
     - トラフィックエンジニアリング
     - トラフィック流入元でのRTBH (Remote Trigger Black Hole)
     - 自AS外でのDDoS Mitigationの実施

- **トラフィックエンジニアリング**
  - 特定リンクの輻輳回避策
    - AS-path prependやMEDを活用して、経路の優先度を下げる
    - 一部のトラフィックを別なリンクへ迂回させ、輻輳の影響を分散させる
    - どうしようもなくなったら特定のリンクへDDoSのトラフィックだけを誘導する覚悟も必要

- **トラフィック流入元でのRTBH (Remote Trigger Black Hole)**
  - IXPでのRTBHの実装が一般的
  - ルートサーバ経由のMultilateral Peering経由の実装が多そう
    - 特定のBlack hole communityを付与した際にnext hopをnullやBH用のnext hopへ書き換える実装
    - ルートサーバで書き換える場合とIXPへ接続しているルータで実装を入れるパターンがある
  - 事前の準備が必要になるので前もって実装を入れておく必要あり
    - IXPごとのRTBHのオプションの仕様把握
    - BH Communityの受け入れ、nexthopを書き換えるroute-mapの設定
    - /24, /48より細かい経路の受け入れ処理

- **自AS外でのDDoSMitigationの実施**
  - DDoS対策のサービスを利用することが一般的
    - 事前にサービス利用の申し込み、Mitigationルールの設定が必要
  - AS外で止めるので相互接続点の輻輳回避が可能
  - ROAの登録に注意
    - Origin ASが異なる、Maximum Lengthの値

これからの相互接続で利用する機能

- **IRR**
  - Hierarchical as-set
- **Convergence Optimization**
  - BFD on a BGP peer
- **BGP Authentication**
  - TCP-AO
- **Routing Security**
  - RPKI ROA / ROV
  - BGP roles
  - ASPA (Autonomous System Provider Authorization)

- **現在、主に利用されているIRR**
  - route/route6
  - as-set

- **as-setの新標準として、Hierarchical as-set (RFC2622) を利用する**
  - 既存のas-setでは名前が被ることがあるため、被らないようにする
  - 今までのas-set名
    - AS-4608 / AS-IIJ
  - Hierarchical as-setの具体例
    - <AS#>:AS-<as-set name>
    - AS17494:AS-CUSTOMERS
    - 名前の衝突解消のため、ASのメンテナーのみ作成可能

```
as-set:         AS174:AS-COGENT
descr:          Cogent and Customers
members:        AS174:AS-COGC-001, AS174:AS-COGC-002, AS174:AS-COGC-003,
                AS174:AS-COGC-004, AS174:AS-COGC-005, AS174:AS-COGC-006,
                AS174:AS-COGC-007, AS174:AS-COGC-008, AS174:AS-COGC-009,
                AS174:AS-COGC-010, AS174:AS-COGC-011, AS174:AS-COGC-012,
                AS174:AS-COGC-013, AS174:AS-COGC-014, AS174:AS-COGC-015,
                AS174:AS-COGC-016, AS174:AS-COGC-017, AS174:AS-COGC-018,
```

```
as-set:         AS2914:AS-GLOBAL
descr:          NTT Global IP Network transit v4 customers
members:        AS2914, AS3949,
                AS2914:AS-US, AS2914:AS-ASIA, AS2914:AS-EUROPE,
                AS2914:AS-SA
```

1. **Hierarchical as-setを作成**
   - Hierarchical as-setの命名ルールでオブジェクトを作成
   - 既存のas-setと同じas-coneをメンバーに入れる
   - この時点では更新が必要であれば、既存のas-setとHierarchical as-setの両方を更新

2. **既存のas-setのメンバーを1で作ったHierarchical as-setに置き換え**
   - 既存のメンバーを更新し、Hierarchical as-setのas-setのみにする
   - 今後はHierarchical as-setのみをアップデートし、既存のas-setはHierarchical as-setに依存させる

3. **PeeringDBをアップデートして対外的に周知**
   - PeeringDBのas-setをHierarchical as-setにアップデート
   - PeeringパートナーやUpstreamプロバイダーへ必要に応じて通知

4. **既存のas-setを削除**
   - 完全にHierarchical as-setへ依存させる
   - 一応、切り戻せるようには準備しておく

- **BFD**
  - Bidirectional Forwarding Detection(BFD)プロトコル
  - ネットワーク内の障害を検出するhelloメカニズム
  - 指定された一定の間隔で hello パケットを送信し、応答がなくなればBGP Neighborを切断する
  - BGPの hello パケットによるkeep alive よりも短い期間で障害を検知し、Black holeで吸い込む事象をBGPによるトラフィックが迂回されることを期待

- **Peerでの利用**
  - IXPのような3rd partyのネットワークを経由するBGPで導入することで品質向上が期待される
    - Hello timerは現状Defaultで運用されていることも多い (180sec / 90secは確かに長い)
    - PeeringDBでは BFD対応済みかどうかのflag機能が実装済
  - ただ、積極的に導入が進められているかというと少し懐疑的
    - 私だけかもしれないが、Peeringを実施する際に特に聞かれていなそう
    - BFDは御存知の通り、片端設定のみでもBGP Peerは確立するので、ネゴられていない？
  - パラメータの設定
    - なやましい、短ければ短いほどよい、というわけでもないよね
    - minimum-interval/minimum-receive-interval: 3000
    - multiplier: 3
    - この例だと 3000msec * 3 で 9000msec (9sec)

- **BGPセキュリティの向上として、MD5key認証の後継となるもの**
- **認証方法**
  - MKT（Master Key Tuple）を定義し、Key material・MAC Algorithms・prieodを管理
  - MKTからtraffic keyを導出し、TCPセグメントごとに認証情報を生成
  - 認証情報をTCP-AOオプションとしてヘッダに付加し、送受信で整合性を検証
- **良い点**
  - アルゴリズムの柔軟性：RFC 5926で複数の暗号方式をサポート、将来的な追加も可能
  - 強固なセキュリティ：MD5より強力な暗号（SHA-1、AES-CMAC、SHA-256など）に対応
  - リプレイ攻撃防御：Sequence Number Extension (SNE)で長期接続を保護
  - 鍵管理の容易さ：1つのセッションへ複数キーを設定可能、セッション確立中にセッション断なしでkeyの切替も可能
- **懸念点**
  - 利用OSに依存：比較的新し目のOSでの実装 (EOS:4.28.2F, IOSXE:17.6.2, IOS XR: 7.0.2, Junos 20.3R1)
  - 設定の複雑さ：MKTやKey Chainの管理が必要
  - CPU負荷増加：強力な暗号化により古い機器では性能低下の可能性
  - 相互運用性課題：MD5とTCP-AOは同時利用不可、両端が対応していないと導入困難

- **TCP MD5とのざっくり比較**

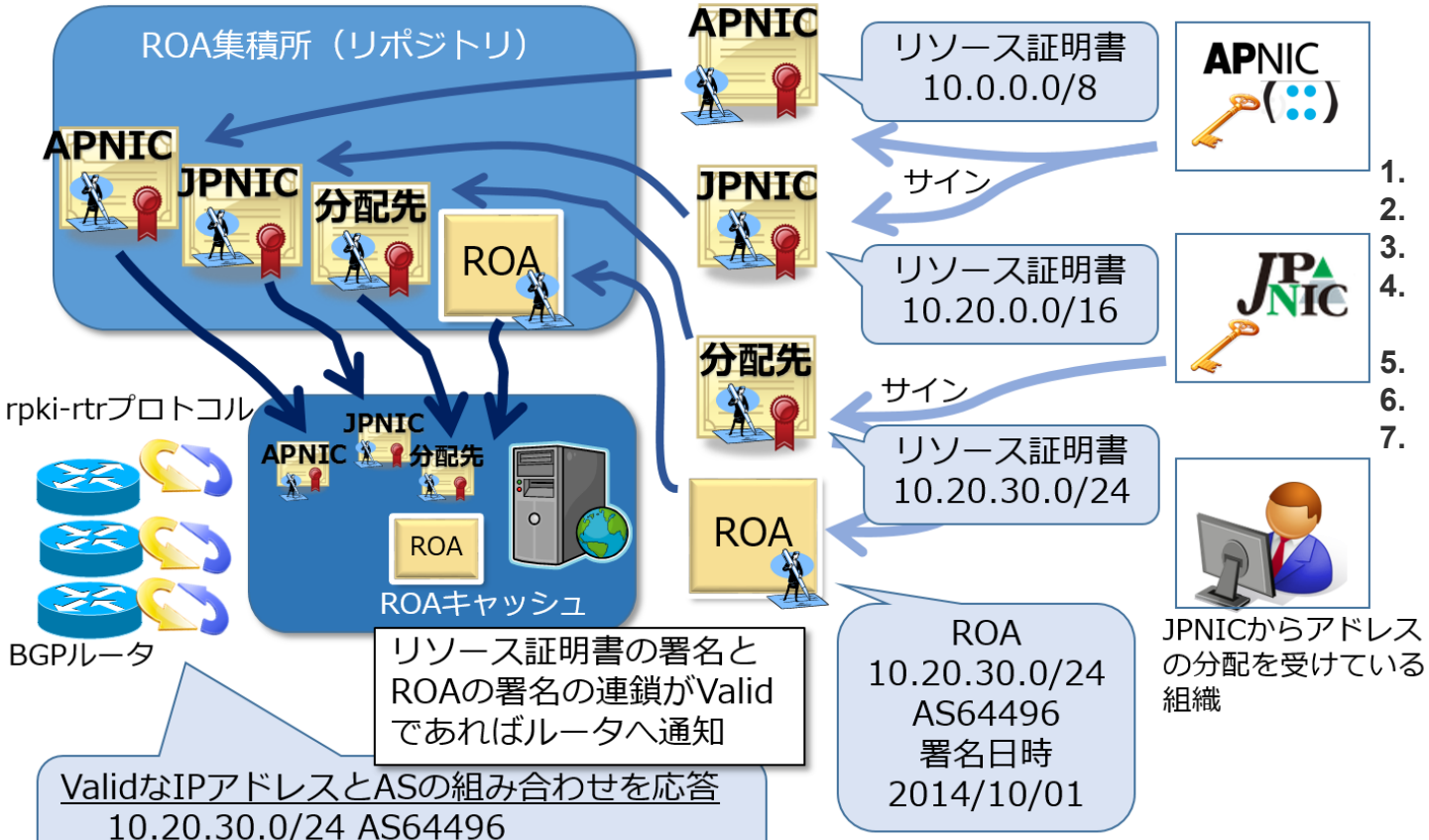
項目	TCP MD5	TCP-AO
RFC	2385	5925/5926
暗号強度	弱い（MD5のみ）	強い（SHA-1, AES-CMAC, SHA-256）
鍵変更	セッション再確立必要	セッションを維持したまま可能
リプレイ攻撃対策	なし	あり
実装普及率	高い	現時点では低い（徐々に増加）
運用負荷	低い	高い（設定複雑/事前ネゴ事項増）

  - MD5は依然として広く使われるが、セキュリティ要件が高い場合はTCP-AOへの移行が有効と考えられる
  - 将来的にはRFC 5925準拠のTCP-AOがPeer/相互接続では標準になるだろう
  - ユースケースとしてIXP上でのPeering, RPKI RTRの暗号化が考えられる

- **TCP-AOを利用するために事前に取り決めておくこと**
  - Master Key（SECRET-DATA）
  - Key Identifier（key-id / send-id / recv-id）
  - MACアルゴリズム
  - 鍵の有効期間（lifetime）

- **導入の目的は経路ハイジャックによるIIJのアドレスの不正利用を<u>インターネット全体</u>で止めること**

1. **保有アドレスのROA(RouteOriginAuthorization)を発行し、外部組織がRPKIによる検証(ROV)を可能とする**
   - 自身が保有するアドレスの経路ハイジャックの抑止を期待
2. **外部組織から受信するBGP経路をRPKI ROV(RouteOriginValidation)で検証し、経路の正当性を担保する**
   - ハイジャックされた経路に誘導しないことでセキュリティリスクや通信不具合から自社のサービスやユーザを守る
   - 自社がインターネットへ広告する経路に不正な経路を含めず、伝搬させないことでインターネット全体の治安を守る



1. JPNICのRPKIシステムへアクセスできる環境を用意
2. アドレス分配のリソース証明書を元にROAを作成
3. RIRが運用するレポジトリへ配置される
4. レポジトリからROAキャッシュ(RelyingPartySoftware)へRRDP/RSYNCで情報を取得
5. TAL(TrustAnchorLocator)によりROA情報を検証し、VRPsを作成
6. 生成したVRPsをrpki-rtrプロトコルによりBGPルータへ配布
7. BGPルータの経路アップデートのタイミングでROVにより経路を検証

出典:https://www.nic.ad.jp/ja/rpki/
•参考:リソースPKIとは(インターネット用語1分解説) : https://www.nic.ad.jp/ja/basics/terms/resource-pki.html

- **ROAの構成要素は3つ**
  - Origin ASN / Prefix / Maximum Length
  - 意図しないROVの発生要因/リスクは上記に含まれる

- **事前検討しておくべき事象**
  - IPアドレスブロックを別なASから広告する
  - IPアドレスブロックの一部を分割して別ネットワークで利用する
  - クラウド型のサービス(Saas)を利用する
  - BYOIP型のサービスを利用する

- **ROAの見直し**
  - 2025年時点でのインターネットにおけるROVの状況
    - IPv4: Valid: 57.18% / Not-Found: 42.68% / Invalid 0.14%
    - IPv6: Valid: 63.32% / Not-Found: 36.25% / Invalid 0.44%
    - Invalidのうち、IPv4 0.07% / IPv6 0.37% が Maximum Lengthの要因
  - 広告経路に対する定期的な見直しが必要
    - 基本的に一度作ったらメンテナンスをしていない、はず

- **BGP Roles**
  - BGPルートリークを、eBGPセッション確立時に役割を合意することにより予防・検出する仕組み
  - オペレータによる経路フィルタ設定ミスや調整漏れに依存しない動作を強制させる
- **概要**
  - **BGP Role Capability**: eBGP確立時のOPENメッセージを拡張し、**Provider/Customer/Peer/Route-Server/RS-Client**の関係を双方で認識、合意する

ロール	役割と経路広告の内容
Provider	リモートASのtransit providerとなるローカルAS。利用可能な経路をCustomerへ広告する
Customer	リモートASのtransit customerとなるローカルAS。ローカルASの経路、ローカルASの顧客経路をProviderへ広告する。それ以外は広告しない
Route-Server	リモートASがRoute Server ClientとなるローカルAS(IXPのルートサーバASを想定)。利用可能な経路をRS-Clientへ広告する
RS-Client	リモートASがRoute ServerとなるローカルAS。ローカルASの経路、ローカルASの顧客経路をRoute-Serverへ広告する
Peer	リモートASもローカルASもピアの関係となるAS。ローカルASの経路、ローカルASの顧客経路をPeerへ広告する。それ以外は広告しない

    - デフォルトの設定では、リモートASからBGP Roleの役割通知がなくとも、ローカルASではBGP Role Capabilityがないことを無視して、BGPセッション確立を進める
    - BGP Role strictモードも存在し、リモートASにBGP Roleの実装を要求する。BGP Roleが合意されなければBGPセッションを確立しない
  - **OTC (Only-To-Customer) Attribute** ：**上流（Provider/Peer/Rute-Server）→下流（Customer）へだけ**伝播すべき経路に付与
  - **OTC Attribute が付いた経路を上流やピアへ再広告するとリーク判定**が可能

- **現時点の状況やユースケース**
  - ベンダー機能実装：各ベンダーでの対応が進んでいる状況
    - BIRD/FRR等のOSS系は既に対応済み
    - Junosは25.2R1から実装対応 (他のベンダーは未確認)
  - ローカルASからの経路リークの自動防止
    - BGP RoleによりBGP Peerの役割を明確化し、上流・ピアでOTC Attributeの経路広告停止を自動化
  - IX/RS環境
    - Route Server/ClientでRoleの合意を使い、OTCを使って経路リークを抑止
  - OTC Attributeの活用
    - OTC Attributeを監視することにより、ルートリークの検知をより迅速に特定および解決できる

- **運用のポイント**
  - 両端実装・合意が必要
    - strictモードでは対向もRFC 9234実装が必要
    - 最初からすべてのBGP Peerで対応することは難しいと予想し、not-negotiatedな状態でのBGP運用を想定しておく
  - OTC Attributeの付与
    - 既存のコミュニティベースのexport/importフィルタとOTC Attributeの整合性を設計する必要あり
    - 基本は混在することを想定。BGP Roleが有効になったところから自動適用を想定する

- **既存のas-setが抱える問題**
  - as-setの中身は再帰的に展開されるが、メンバー数や再帰の深さに制限がなく、無制限に膨張する可能性がある
  - 内容の品質や整合性、真正性を保証する仕組みがない

- **実例：2024年3月のBGPリーク**
  - ロシアのMTS（AS8359）が、香港IX（AS4635）から3万以上の経路を誤ってグローバルにリーク
  - as-set「AS-MTU」が使われており、その中身は4万以上のASNを含み、グローバルのASN総数（約8.3万）に匹敵
  - このような巨大as-setをフィルタとして使うと、膨大な経路リストが生成され、ルーターの設定が非現実的な規模になる

- **解決先を模索中**
  - as-setの所有者が内容を最小限に見直し、再帰の深さも抑制することが推奨されるが、グローバルな協力が必要で現実的には困難
  - 最終的には、IRRベースのas-setに頼らずRPKIベースのASPA検証で受け入れPrefixを制御する必要がある

- **RPKIベースの ASPA**
  - RPKI ASPAの仕組みで、受信したAS_PATHが顧客→プロバイダの関係に沿っているか検証する仕組み
  - ASごとに一意のASPA署名オブジェクトを作成し、自ASの「上流プロバイダ情報」を記載
    - ASx, {ASxxx, ASyyy, ASzzz}
  - 既存のRPKIの仕組みを利用し、RIR/NIRのレポジトリへASPAオブジェクトを配置
  - RPKIキャッシュサーバでASPAオブジェクトを取得し、AS_PATH検証用のVAPsを作成
  - RPKIキャッシュサーバからRTRでVAPsを共有
    - このあたりはRPKI ROA/ROVと同じような動き
  - 現状IETFでDraft議論が継続、標準化進行中

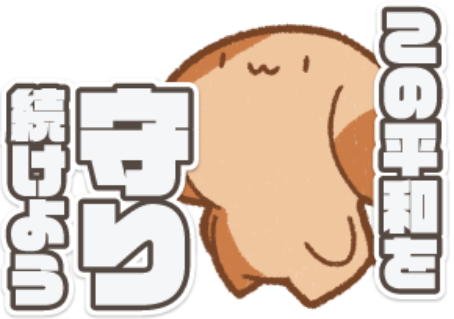
- **BGP Roleで送信側制御＋RPKI ASPAで受信側検証による二重防御が想定される**
  - ローカルASからの経路リークを防御 : BGP Roleによる経路リークの自動制御
  - リモートASからの経路リークを防御 : ASPAによる不正経路検証による流入停止

- **今後の未来予想**

	現在	ちょっと先の未来	まだ先の未来
Route Object	• 既存のas-setをHierarchical as-setに置き換え • as-set/IRRの登録内容の棚卸し/メンテナンスを頑張る (年1回くらい)	• Hierarchical as-set と IRR を維持、メンテナンス	• Hierarchical as-set と IRR を維持、メンテナンスが不要になる
Convergence Optimization	• BFDを活用を検討	• BFDを活用	• BFDを活用
BGP authentication	• MD5をBGPネイバーごとに交渉して設定	• TCP-AOの接続認証とMD5をBGPネイバーごとに並行運用 • TCP-AOが実施できるかの確認と内容のすり合わせを行う	• TCP-AOによる接続認証を行う

Routing Security	• RPKI ROA/ROV
	  • ROAは必ず登録
	  • ROVは適切に運用
	• import/exportの経路制御ポリシーは適切に設定
		• RPKI ROA/ROV が継続
		• import/exportの経路制御ポリシーは適切に設定
		• BGP Roleによる経路広告制御を並行運用
			• RPKI ASPAオブジェクトを登録
			• RPKI ROA/ROV によるOrigin AS、RPKI ASPAによるAS-Path検証
			• BGP Roleによる経路広告制御
			• import/exportの経路制御ポリシーは必要に応じて実施

- **未来は誰もわからないけど、インターネットは互助・共栄の世界**
- **みんなでこの平和を守り続けよう**



日本のインターネットは1992年、IIJとともにはじまりました。以来、IIJグループはネットワーク社会の基盤をつくり、技術力でその発展を支えてきました。インターネットの未来を想い、新たなイノベーションに挑戦し続けていく。それは、つねに先駆者としてインターネットの可能性を切り拓いてきたIIJの、これからも変わることのない姿勢です。IIJの真ん中のIはイニシアティブ――IIJはいつもはじまりであり、未来です。