

# P2MP TEを用いたVPNマルチキャスト ネットワークについて

---

NTT

安川正祥

([yasukawa.seisho@lab.ntt.co.jp](mailto:yasukawa.seisho@lab.ntt.co.jp))

# 発表の内容

---

- MVPN方式まとめ
  - VPNサービス要求条件
  - VPNネットワーク要求条件
  - 方式詳細とその適用領域
  - MVPNシナリオ (P2MP TE & BGPシナリオ含む)
  - シナリオ評価
  - 課題
  - まとめ
-

# MVPNプロトコル

## draft-ietf-l3vpn-2547bis-mcast-00

PE-CE		PE-PE						
CNTL plane	<ul style="list-style-type: none"> <li>*PIM-SM</li> <li>*PIM-SSM</li> <li>*PIM-Bidir</li> <li>*PIM-DM</li> </ul>	Autodiscovery of MVPN membership	PE-PE Transmission of Multicast Routing info	Aggregation	P-multicast tree-MVPN mapping	Switching to S-PMSIs	Inter-AS Procedures	C-RPs co-location on a PE
		<ul style="list-style-type: none"> <li>*BGP-based AutoDiscovery (PEAddr, RD, RT, RPFAttribute, I-PMSI Attribute, ID for tunnel Encapsulation type, Aggregation McastCNTLtransmission scheme)</li> <li>*PIM and MI-PMSI based AutoDiscovery</li> </ul>	<ul style="list-style-type: none"> <li>*PIM Peering Full Per-MVPN PIM peering across a MI-PMSI(rosen08)</li> <li>-Lightweight PIM peering across a MI-PMSI (Hello omission RefreshReduction)</li> <li>Unicasting of PIM C messages</li> <li>*BGP carrying Multicast routing A new SAFI, RD, PEAddr, C-Src C-Grp, upstream PEAddr, RT</li> </ul>	<ul style="list-style-type: none"> <li>*Aggregation</li> <li>*No Aggregation</li> </ul>	<ul style="list-style-type: none"> <li>*Configuration BGP based advertisement after root's leaves discovery</li> <li>*BGP based Advertisement As part of MVPN membership discovery</li> </ul>	<ul style="list-style-type: none"> <li>*UDP-based protocol (for MI-PMSI, S-PMSI Join)</li> <li>*BGP-based protocol (for UI-PMSI &amp; MI-PMSI)</li> <li>*Two methods for ingress to switch data from I-PMSI to S-PMSI</li> </ul>	<ul style="list-style-type: none"> <li>*PMSI based on a tunnel spanning multiple ASs</li> <li>*PMSI based on independent tunnels for ASs</li> </ul>	<ul style="list-style-type: none"> <li>*C-RPs co-location on a PE</li> <li>*No C-RPs co-location on a PE</li> </ul>
FWD plane	*IP(Mcast)	P-Multicast Service Interface type	Tunnels instantiating PMSIs type		Default method of transmitting multicast data		Encapsulation	
		<ul style="list-style-type: none"> <li>*MI-PMSI</li> <li>*UI-PMSI</li> <li>*S-PMSI</li> <li>*Aggregate MI-PMSI</li> <li>*Aggregate UI-PMSI</li> <li>*Aggregate S-PMSI</li> </ul>	<ul style="list-style-type: none"> <li>*PIM Trees (PIM-SSM, PIM-Bidir, PIM-SM)</li> <li>*RSVP-TE P2MP</li> <li>*Meshed Unicast Tunnels</li> <li>*Unicast Tunnels to the Root of a P-Tree</li> <li>*LDP Trees</li> </ul>		MI-PMSI or UI-PMSI		<ul style="list-style-type: none"> <li>* GRE</li> <li>* MPLS</li> <li>* IP-in-IP</li> </ul>	

# どうやってこれらのオプションを使いこなすの？

- ・BGPとPIMベースのどちらのオプションを使うの？
- ・[U/M]I-PMSI、S-PMSI、Aggregationの関係は？
- ・PIM GRE、mLDP、P2MP TEのうちどのトンネルを利用するの？
- ・RP-colocation は?...

- ・もう既に Rosen-MVPNを導入済みだけどこれらの新しい機能使えるの？
- ・どんなマイグレーションシナリオが期待できるの？
- ・ルータ丸ごと変更する必要があるの？  
それともアップグレード可能？
- ・IP unicast VPNへの影響は？

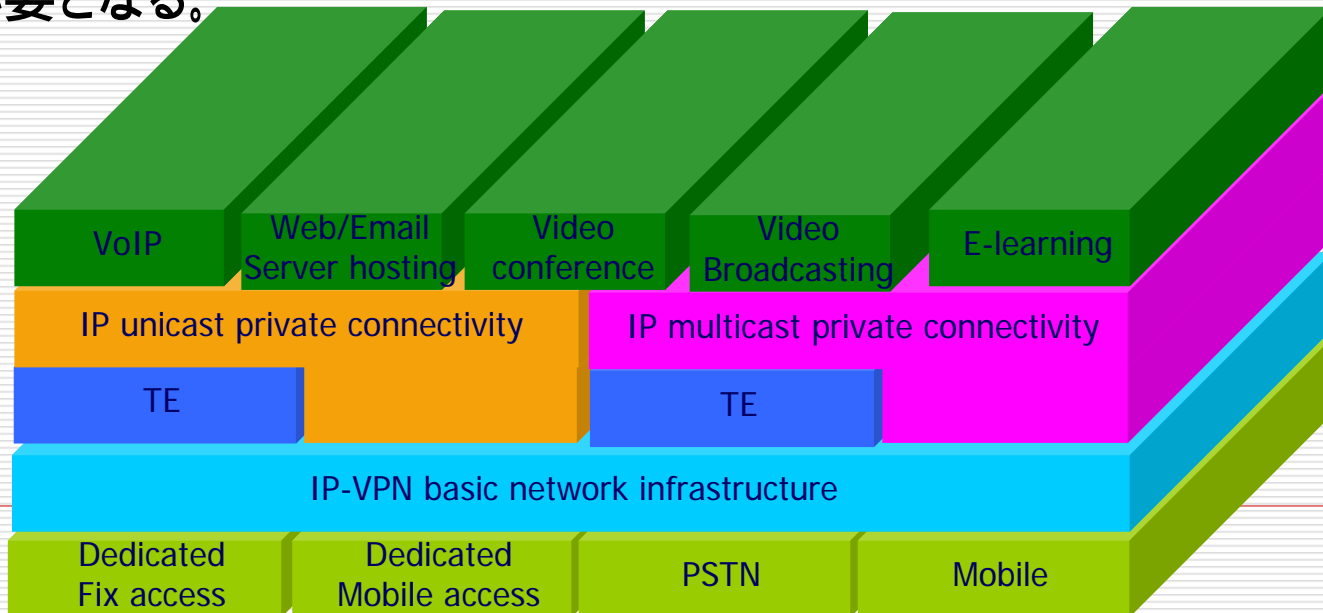


- ・異なる要求条件の色々なマルチキャストVPNサービスをどのように展開するの？
- ・MVPNのスケーラビリティ性能をどのようにして向上させるの？
- ・QoSを保証したプレミアムマルチキャストVPNサービスは実現できる？

- ・このMVPNをどうやって管理する？
- ・PMSIのトンネルはどこに設定されているの？
- ・ユーザのマルチキャストVPNサービスの品質はちゃんと担保できている？
- ・故障・障害からどのようなアクションをとればよいの？

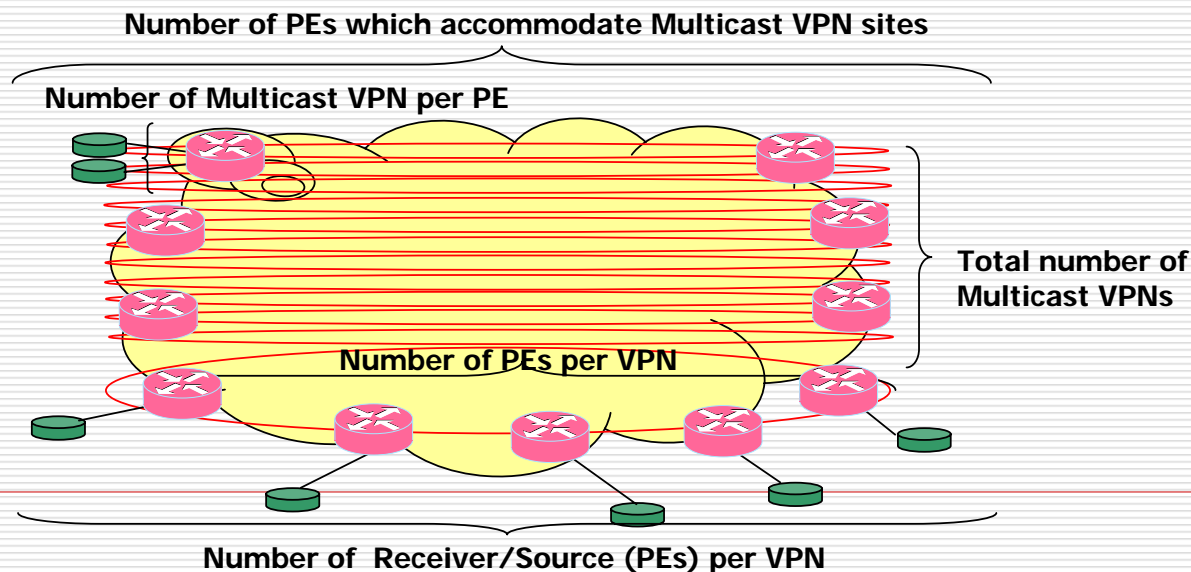
# マルチキャストVPNサービス要求条件

- 多種多様のマルチキャストVPNサービスをIP-VPN上に実現。
- ブロードバンドビデオ放送、ビデオ会議、IPラジオ、Eラーニング、ライブビデオ配信、金融データ配信、POSデータ配信 ...
- マルチキャストVPNサービスはVPNアーキテクチャに多大な影響(帯域、対置数が大きい)を及ぼす。
- サービスプロバイダには柔軟でスケール性の高いアーキテクチャが必要となる。



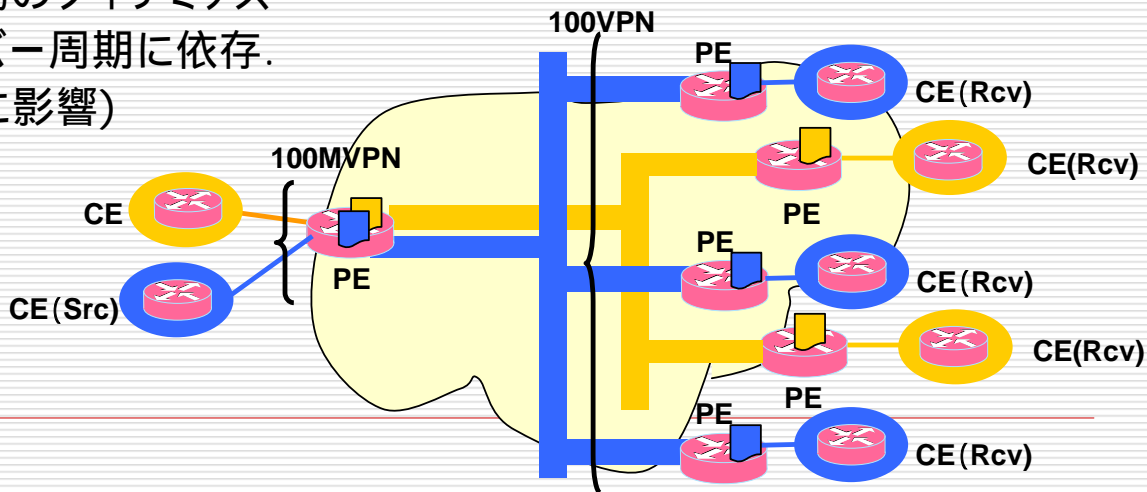
# Multicast VPNスケール規模

- ❑ Multicast VPN数: ~数千
- ❑ PE当たりのMulticast VPN数: ~数百
- ❑ VPN当たりのPEs数: ~数百
- ❑ Multicast VPNサイトを収容するPE数: ~数百
- ❑ PE当たりのMulticast group数: ~数千
- ❑ VPN当たりのMulticast group数: ~数十



# Multicast VPNのダイナミクス

- ❑ Multicast VPNサービス起動時のダイナミクス  
(e.g. 1000VPNx10group > 30%起動 > 3000group > 100 PE-PE > 300,000 PIM Hello/Join/Prune起動.)
- ❑ PEにおけるC-PIMインスタンスの保持のダイナミクス  
(e.g. 100MVPN/PE, 100PE/VPN > 10,000 PIM neighbor/PE > 330Hello msgs/sec)
- ❑ I-PMSIトンネル起動のダイナミクス  
(VPNサイト追加/削除周期に依存. Autodiscovery時のBGP updateに影響)
- ❑ S-PMSIトンネル起動のダイナミクス  
(S-PMSIスイッチオーバー周期に依存. スwitchオーバー広告に影響)



# Multicast VPNマネジメント要求条件

---

- Multicast VPNサービスではより高度のSLAが要求される
  - マルチキャストグループ全体の品質確保
  - マルチキャストグループ全体にわたる故障特定・回復
- Multicast VPN ネットワークではより高度なネットワーク運用が要求される
  - 複数のソース/リーフ間のマルチキャストコミュニケーション状態のモニタリング
  - 複雑なトラブルシューティング、部分リーフのマルチキャスト通信の障害回復など

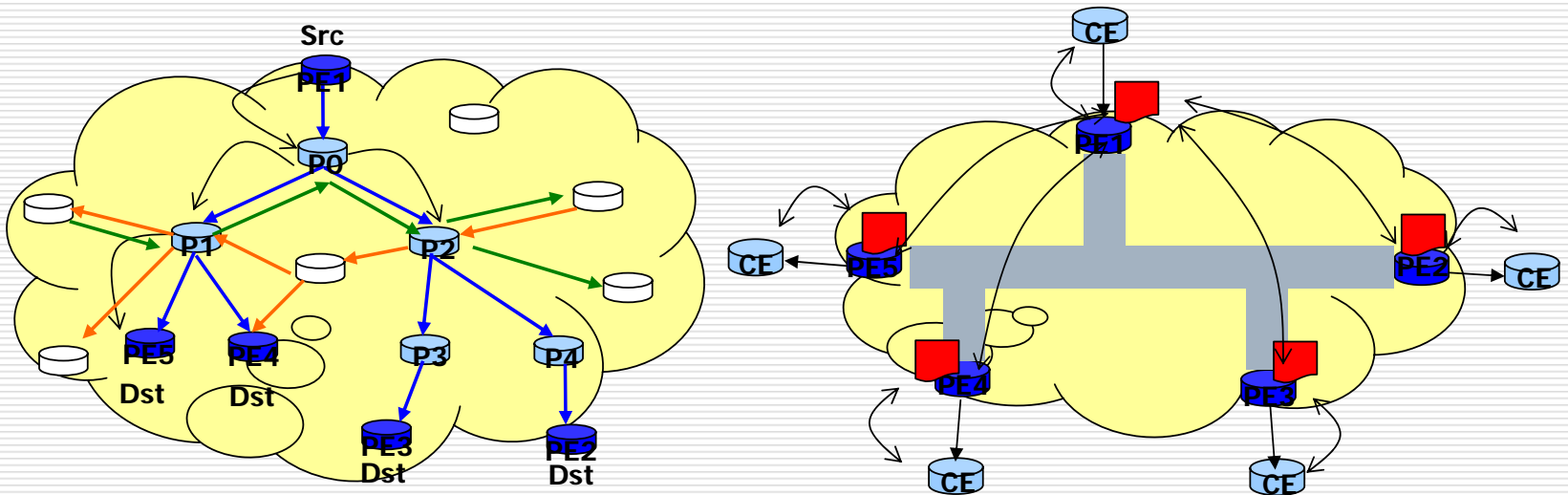
経路がツリーかされ通信対置が増えるだけで大変。。。。

---



# PIM multicastツリーのモニタリング

- ❑ PIMツリーの経路と経路上の通信状態の把握が必要
- ❑ マルチキャスト関連情報を取得するのに"show ip mroute", "show ip pim neighbor"などのコマンドが利用できるが。。。
- ❑ マルチキャストOAMツール、"mtrace"などの完成度、使い勝手は??.



# MVPN membershipの発見・保持機構

- MVPN membershipの発見・保持機構
    - BGPベースのAutodiscovery
    - PIMとInclusive treeベースのAutdiscovery
  - P-core上でのPIM Hello交換削除によるMVPN CNTL planeのスケラビリティ性能向上
  - 複数のPMSIsとマルチキャストトンネルを利用するためにはVPNメンバ間で多くの能力情報の交換が必要
  - IPユニキャストVPNでは既にBGPベースのAutodiscovery機構が利用されている
  - BGPベースのAutodiscovery機構を採用
-

# PE-PE間のマルチキャストルーティング情報の交換機構

---

- PE-PE間のマルチキャストルーティング情報交換
    - MI-PMSI上の完全PIMピアリング
    - MI-PMSI上の簡易PIMピアリング(Hello削減、PIM RR)
    - PIM C-Join/Prune msgsユニキャスト
    - BGPによるPE-PE間マルチキャストルーティング情報交換
  - 周期的なPIM制御メッセージ交換の削減によりMVPNの制御プレーンのスケール性能を向上
  - 簡易PIMもしくはBGPによる情報交換を推奨
  - ルートへ接続しない無関係なPEへの不要なC-JOIN/PRUNE msg転送を防止
  - 使用するPMSIに依存しないPE-PE間の情報交換機構を推奨
-

# PMSIsの選定

---

- P-Multicast Service Interface (PMSI)
    - MI-PMSI, UI-PMSI, S-PMSI
    - Aggregate MI-PMSI, UI-PMSI, UI-PMSI
  - デフォルトのP-Multicast Serviceに対してはMI-PMSIもしくは Full meshのI-PMSIを使用.
  - P-Multicast Service転送の最適化にS-PMSIを利用
  - 2つのメジャーなPMSIの組み合わせ.
    - 1) MI-PMSI & S-PMSI (PIMベースのアプローチ)
      - MI-PMSI: C-message & Multicastデータ
      - S-PMSI: Multicast dataデータ
    - 2) UI-PMSI & S-PMSI (BGPベースのアプローチ)
      - UI-PMSI: Multicastデータ (スタティック)
      - S-PMSI: Multicastデータ (ダイナミック)
-

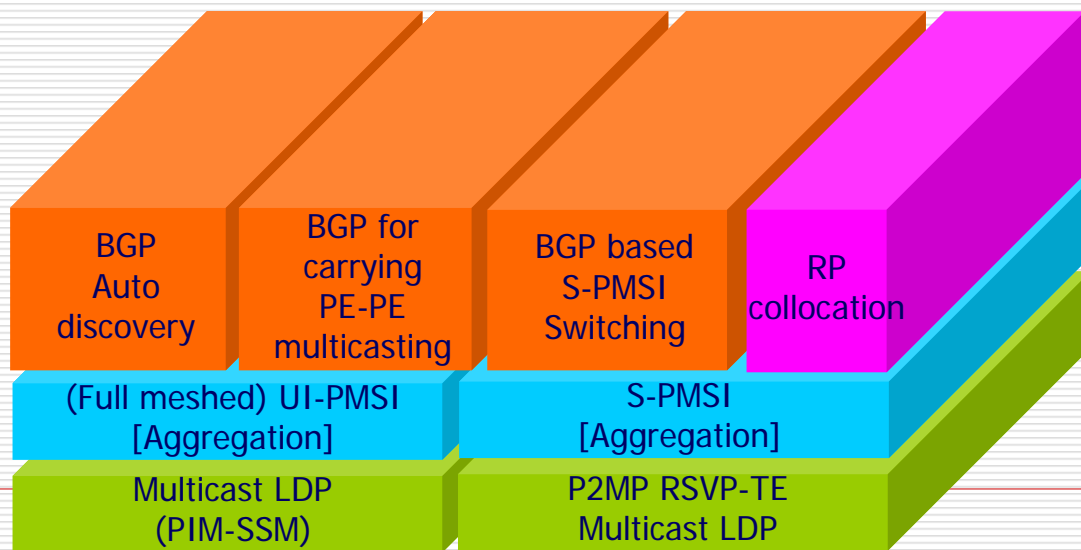
# PMSIのトンネル技術

---

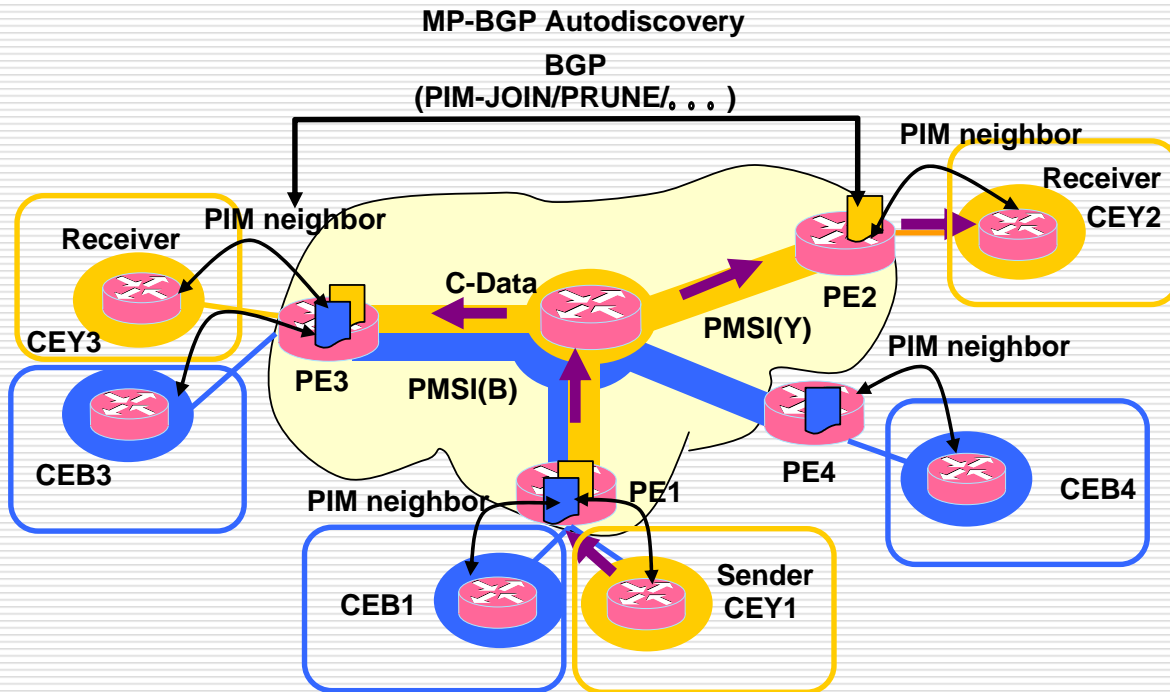
- P-Multicast Service Interface.
    - PIM: PIM-BIDR, PIM-SM, PIM-SSM
    - MPLS: P2MP RSVP-TE, Multicast LDP, Unicast MPLS
  - PMSIトンネル技術
    - MI-PMSI: PIM-BIDR, PIM-SM, Multicast LDP\*
    - UI-PMSI: PIM-SSM, P2MP RSVP-TE, Unicast tunnel, Multicast LDP
    - S-PMIS: PIM-SSM, P2MP RSVP-TE, Unicast tunnel, Multicast LDP
  - RosenではGRE PIMが利用されており、PIM-SSMがメジャー
  - トンネル技術のMPLS統合がルータの転送性能の向上に寄与
  - P2MP RSVP-TEトンネルをS-PMSIに用いることで高信頼のQoS保証型のプレミアムVPNマルチキャストサービスが実現可能
-

# 1. BGP + MPLSシナリオ

- MVPNの制御プレーンをBGPに転送プレーンをMPLSに統一するシナリオ
- デフォルトのVPNマルチキャストトラフィック転送としてLDPマルチキャストを用いたフルメッシュのUI-PMSIを使用し、高品質・信頼のVPNマルチキャストトラフィック転送としてP2MP RSVP-TEを用いたS-PMSIを使用
- VPNユーザのプロバイダNW内のマルチキャスト転送を最適化するためにPEへのRP共存メカニズム、MVPNトラフィック間のアグリゲーションを利用

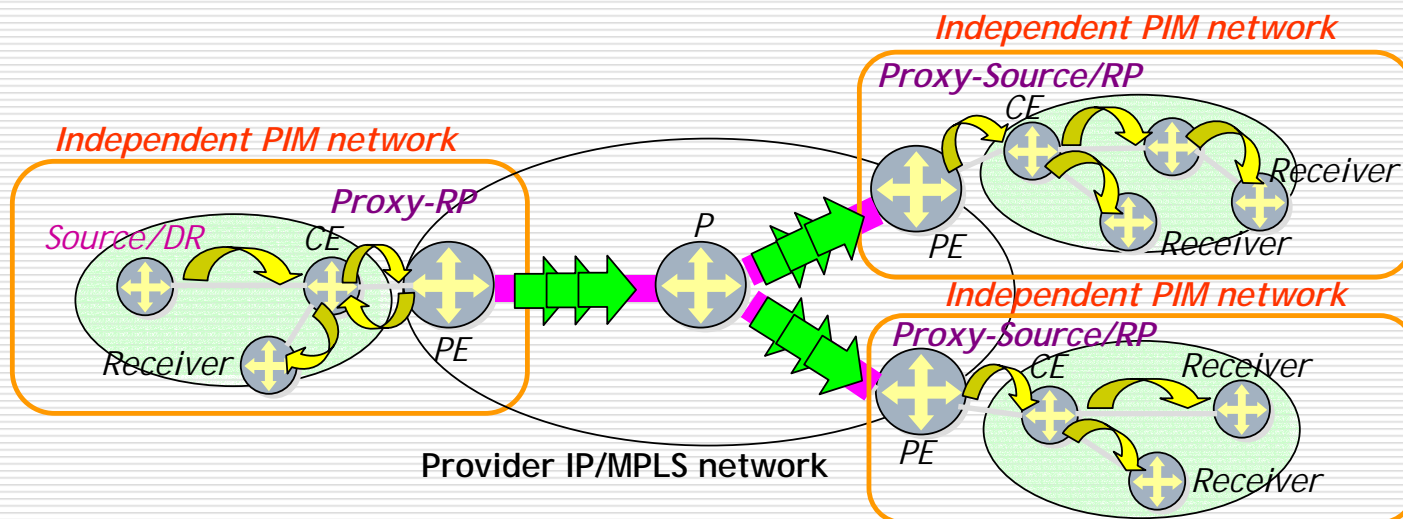


# ネットワークアーキテクチャ概要



# Proxy-Source/RP model

- ❑ VPNのメンバである全てのPEがproxy-Source/RPsとして動作.
- ❑ Proxy-Source/RPとして動作するPEはVPNユーザのPIM msgsを終端, これによりサイト内の独立なPIM運営が可能.
- ❑ PE間をP2MP TEトンネルで相互接続しマルチキャストトラフィックを転送.



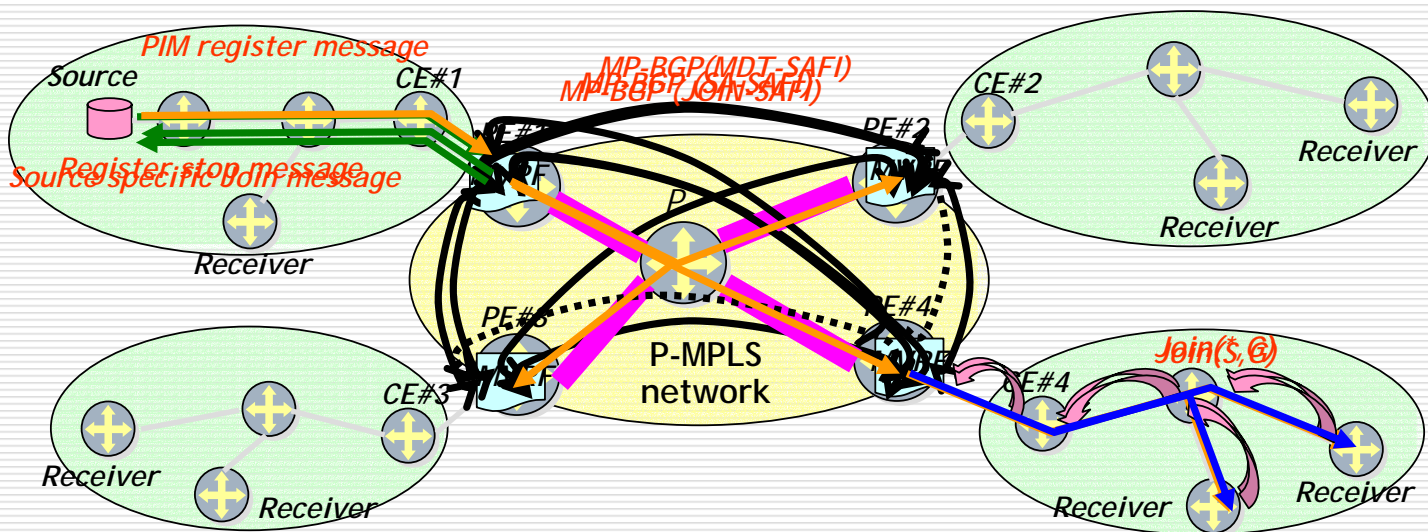


# PIM-SM operation example

Configuring MVRFs on PE triggers exchange of MDT-SAFI information

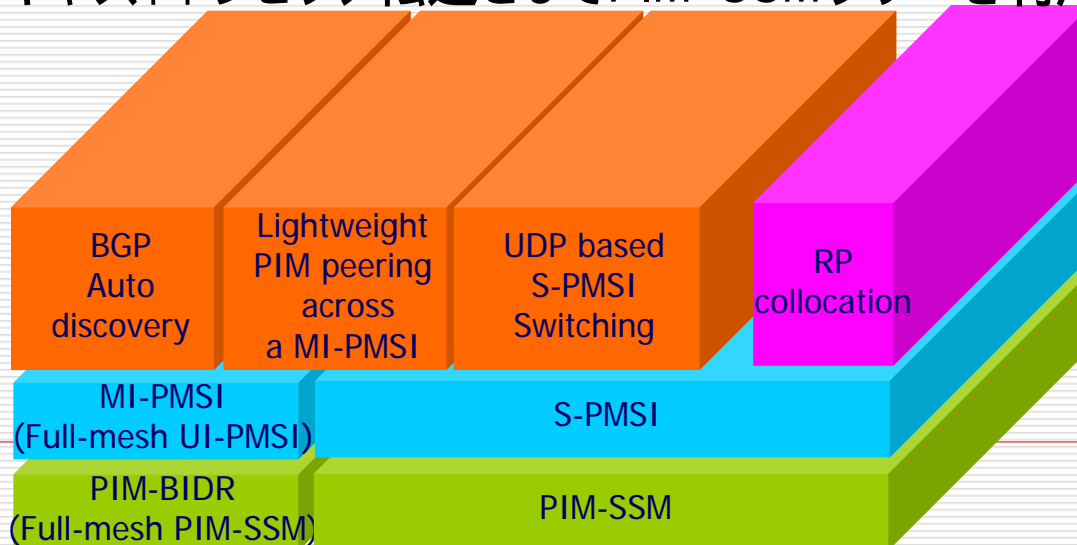
Interested receivers send their joins to the PEs (the RP for this site)

Default MDT: Each PE sets up P2MP tunnel to other PEs of the same VPN



## 2. PIM + GREシナリオ

- VPN制御プレーンとしてVPNメンバシップのauto-discoveryにBGPを用いて、VPNマルチキャストの制御メッセージ交換にPIMを利用し、転送プレーンをGREに統一するシナリオ。
- PIMベースの制御プレーンを構成するために、MI-PMSIもしくはフルメッシュのUI-PMSIを設定し、プロバイダコアネットワークを跨ったPIMpeeringによるPIM制御メッセージの交換を行う。
- デフォルトのVPNマルチキャストトラフィックの転送としてフルメッシュのPIM-SSMツリーもしくはPIM-BIDRツリーを用いて、高ボリュームのVPNマルチキャストトラフィック転送としてPIM-SSMツリーを利用する。



# 課題

---

- Rosen MVPNから新しいIMVPNへのマイグレーションシナリオが必要.
- Multicast VPNネットワーク監視・管理ツールの整備.
- 新しいプロトコルの評価
  - P2MP LDP
  - 簡易PIM (Hello削減, PIM RefreshReduction)
  - P2MP RSVP-TE (Inter Areaなど)

# まとめ

---

- Multicast VPNサービス要件をまとめた.
- Multicast VPNネットワークシナリオとして2つの大きな方向性があることを紹介.
  - BGP+MPLS
  - PIM+GRE
- P2MP RSVP-TEを用いたVPNマルチキャストアーキテクチャについて紹介.