

## BGP/MPLS VPNs

Yakov Rekhter  
yakov@juniper.net

## Agenda

- ➔ ■ BGP/MPLS MVPN – what are the goals ?
- Supporting PIM-SM in SSM mode MVPNs
- Supporting PIM-SM in ASM mode MVPNs
- Summary

## draft-rosen – control plane problems

- **Control plane overhead:**
  - Each PE has to maintain PIM adjacencies with all other PEs for which it has at least one MVPN in common
    - If  $M$  PEs have  $N$  MVPNs in common, then each such PE maintains  $M*N$  PIM adjacencies with the other PE
    - Due to the virtual router model
  - Further compounded by complete period update property of PIM
    - Every 60 secs
- **Control plane limitations:**
  - Inter-AS/inter-provider operations options (b) or (c) require PEs in different ASes/providers to have (direct) PIM routing peering with each other
    - As long as these PEs have at least one MVPN in common
    - Due to the virtual router model

## draft-rosen – data plane problems

- **Data plane overhead:**
  - No ability to aggregate multiple MVPNs into a single inter-PE tunnel
- **Data plane limitations:**
  - No MPLS support
  - Forces all providers to use the same tunneling technology - GRE

## BGP/MPLS MVPN – what are the goals ?

- **Extend 2547 VPN service offering to include support for IP multicast for 2547 VPN customers**
- **Follow the same architecture/model as 2547 VPN unicast**
  - No need to have the Virtual Router model for multicast and the 2547 model for unicast
- **Re-use 2547 VPN unicast mechanisms, with extensions, as necessary**
  - No need to have PIM/GRE for multicast and BGP/MPLS for unicast
- **Retain as much as possible the flexibility and scalability of 2547 VPN unicast**

## Agenda

- BGP/MPLS MVPN – what are the goals ?
- ➔ ▪ Supporting PIM-SM in SSM mode MVPNs
- Supporting PIM-SM in ASM mode MVPNs
- Summary

## IP Multicast with PIM-SM in SSM mode and MVPN

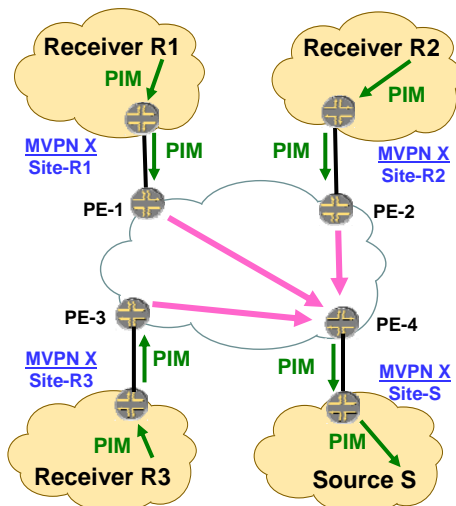
### Plain IP Multicast:

- Multicast Sources need to know that there are Multicast Receivers – the Receivers have to inform the Sources that the Receivers want to receive traffic from the Sources
- With PIM-SM in SSM mode the Receivers discover the Sources by means outside of PIM
- There has to be multicast forwarding state from the Sources to the Receivers to carry multicast traffic from the Sources to the Receivers

### In the context of MVPN:

- Carrying multicast routing information from the Receivers to the Sources may involve MVPN service providers
  - As multicast sources and multicast receivers may be in different sites
- Carrying multicast data traffic from the Sources to the Receivers may involve MVPN service providers
  - As multicast sources and multicast receivers may be in different sites
- Different MVPNs may use the same address space (e.g., RFC1918), including IP multicast addressing space

## Carrying MVPN multicast routing information (from the Receivers to the Source)



**Step 1:** from the Receivers to the PEs of the sites that contain the Receivers (PEs of Site-R1, Site-R2, Site-R3)

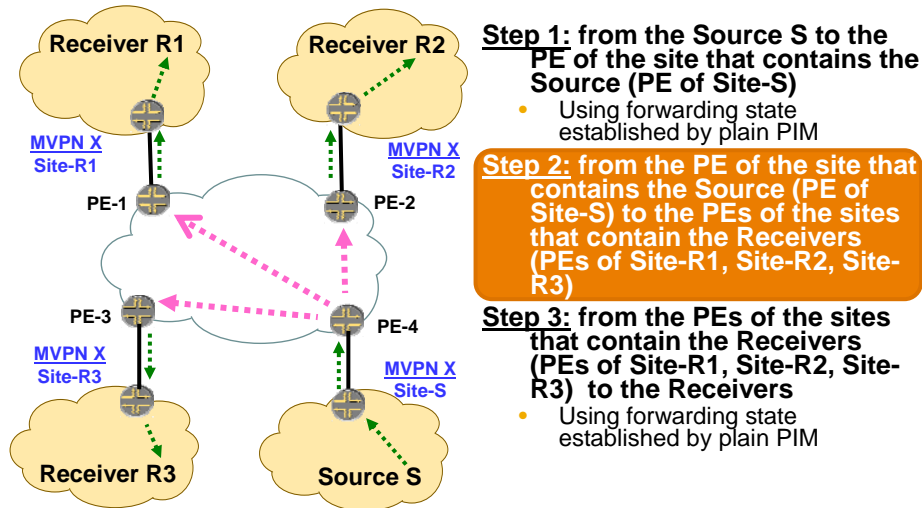
- Using plain PIM

**Step 2:** from the PEs of the sites that contain the Receivers (PEs of Site-R1, Site-R2, Site-R3) to the PE of the site that contains the Source (PE of Site-S)

**Step 3:** from the PE of the site that contains the Source (PE of Site-S) to the Source S

- Using plain PIM

## Carrying MVPN multicast data traffic (from the Source to the Receivers)



## MVPN – what are the required service provider mechanisms ?

- A mechanism to carry MVPN multicast routing information from the PEs connected to the sites that contain the Receivers to the PE connected to the site that contains the Source**
  - E.g., from the PEs connected to Site-R1, Site-R2, Site-R3 to the PE connected to Site-S
  - So that Receivers inform the source S that the Receivers want to receive traffic from S
- A mechanism to carry multicast traffic from the PE connected to the site that contains the Source to the PEs connected to the sites that contain the Receivers**
  - E.g., from the PE connected to Site-S to the PEs connected to Site-R1, Site-R2, Site-R3
  - So that multicast traffic will flow from the source S to the Receivers

## Agenda

- BGP/MPLS MVPN – what are the goals ?
- Supporting PIM-SM in SSM mode MVPNs
- ➔ • Carrying MVPN multicast routing information
  - Carrying MVPN multicast traffic
- Supporting PIM-SM in ASM mode MVPNs
- Summary



## Carrying MVPN multicast routing information: BGP C-multicast routes

- **BGP C-multicast (customer multicast) routes carry MVPN customer multicast routing information from the PEs connected to the sites that contain the Receivers to the PE connected to the site that contains the Source**
- **C-multicast routes are carried in Multiprotocol BGP (RFC4760) using new NLRI – MCAST-VPN**
- **C-multicast route NLRI contains:**
  - Multicast Source (S), Multicast Group (G)
  - Route Distinguisher (RD)
    - Needed to support MVPNs that may use the same address space (just like with unicast)
- **Import of a C-multicast route into a VRF is controlled by the Route Target carried by the route**
  - A C-multicast route is imported into a VRF if the VRF Route Import of this VRF matches the Route Target carried in the C-multicast route
- **Re-use the existing BGP mechanisms (e.g., extended communities, Route Target constraint, Route Reflectors, etc..)**

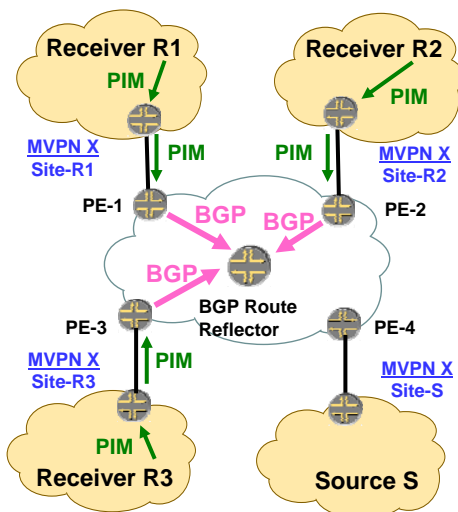
## Import of C-multicast routes

- **C-multicast route carrying a given Multicast Source S should be imported ONLY into the VRF on the PE connected to the site that contains S**
  - And not into any other VRF, even of the same MVPN
- **To accomplish this each VRF on a given PE has a distinct VRF Route Import extended community associated with it:**
  - Contains PE's IP address + local (to the PE) number ->
    - Different MVPNs within a given PE have different Route Imports
    - Within a given MVPN VRFs on different PEs have different Route Imports
  - VRF Route Import is auto-configured

### AND

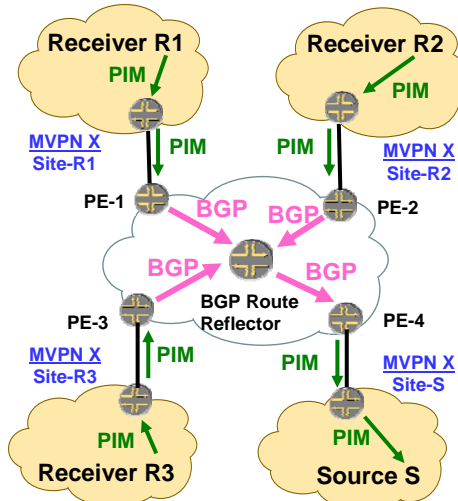
- **All the VRFs within a given MVPN have the information about VRF Route Imports of each of these VRFs**
  - Accomplished by piggybacking VRF Route Import extended community on the (unicast) VPN-IPv4 routes
- **To make a C-multicast route carrying Multicast Source S be imported only into the VRF on the PE connected to the site that contains S:**
  - find the (unicast) VPN-IPv4 route to S
  - set the Route Target of the C-multicast route to the VRF Route Import carried by the found VPN-IPv4 route

## Originating C-multicast route



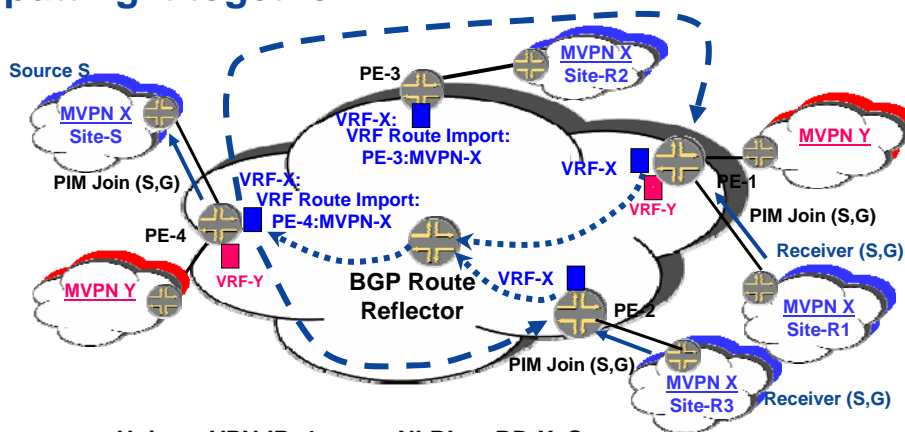
- **PE-1 determines that Site-R1 has receiver(s) for (S,G)**
  - PE-1 receives from Site-R1's CE PIM Join (S, G)
- **PE-1 constructs and originates a C-multicast route as follows:**
  - Finds (unicast) VPN-IPv4 route to S
  - Extracts from this route: RD and VRF Route Import
  - C-multicast route carries:
    - <Source, Group> - from PIM Join (S, G)
    - RD - from the VPN-IPv4 route
    - Route Target - constructed from VRF Route Import of the VPN-IPv4 route
- **Same applies to PE-2 and PE-3**

## Receiving C-multicast route



- PE-4 receives the C-multicast route originated by PE-1 (PE-2, PE-3)
  - These C-multicast routes are aggregated by BGP Route Reflector
- PE-4 accepts the C-multicast route into the VRF for MVPN X
  - Because the VRF Route Import on this VRF matches the Route Target carried in the C-multicast route
- PE-4 creates (S,G) state in the VRF, and propagates <S,G> information to the CE of Site-S (the site that contains S)
  - Using PIM as PE-CE multicast routing protocol

## Carrying MVPN multicast routing information – putting it together



- ← — Unicast VPN-IPv4 route: NLRI = <RD-X, S>, VRF Route Import = <PE-4:MVPN-X>
- ← ... C-multicast route: NLRI = <RD-X, S, G>, Route Target = <PE-4:MVPN-X>



## Agenda

- BGP/MPLS MVPN – what are the goals ?
- Supporting PIM-SM in SSM mode MVPNs
  - Carrying MVPN multicast routing information
- ➔ • Carrying MVPN multicast traffic
- Supporting PIM-SM in ASM mode MVPNs
- Summary



## Carrying MVPN multicast traffic: inter-PE tunnels and BGP auto-discovery routes

- Inter-PE tunnels are used to carry MVPN multicast data traffic from the PEs connected to the sites that contain the Sources to the PEs connected to the sites that contain the Receivers
- **BGP auto-discovery routes perform two functions:**
  - (1) Enable establishment of inter-PE tunnels
    - Auto-discovery routes do NOT (directly) establish tunnels – tunnels are established by the appropriate signaling protocol associated with a particular type of a tunnel
    - Signaling protocols use the information carried in the auto-discovery routes
  - (2) Bind one or more MVPN multicast <Source (S), Group (G)> streams to a particular inter-PE tunnel
    - Many-to-one binding

## More on auto-discovery routes

- **Auto-discovery routes are carried in Multiprotocol BGP (RFC4760) using MCAST-VPN NLRI**
- **Handled similar to VPN-IPv4 routes:**
  - RDs to distinguish among different MVPNs
  - import/export based on Route Target extended communities
- **Re-use the existing BGP mechanisms (e.g., extended communities, Route Target constraint, Route Reflectors, etc...)**

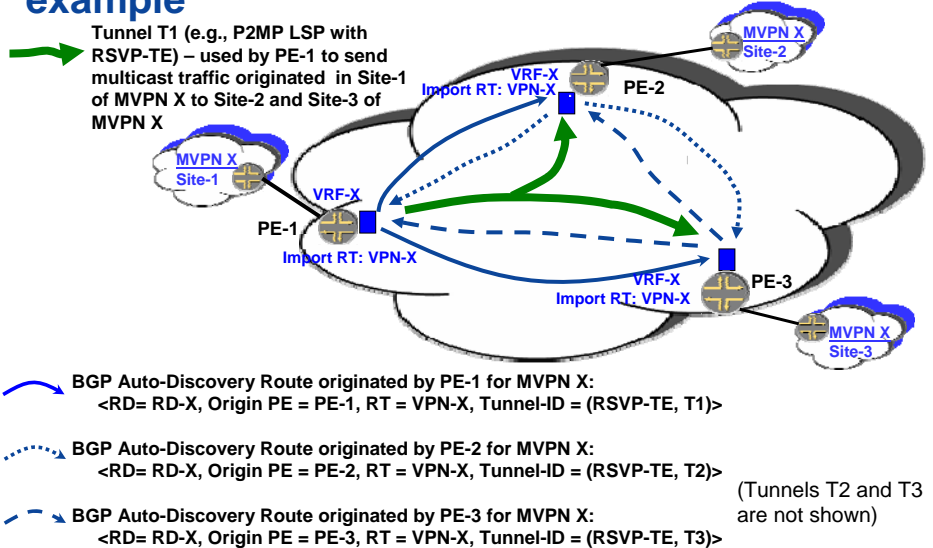
## More on auto-discovery routes (cont.)

- **PE connected to a MVPN site advertises an auto-discovery route that carries the export Route Targets for that MVPN configured on the PE**
- **PE connected to an MVPN site imports received auto-discovery route into the VRF of that MVPN only if at least one of the import Route Targets configured on the PE for that MVPN matches the Route Target carried in the auto-discovery route**
- **MVPN import/export Route Targets can be the same as Route Targets used by unicast 2547 VPNs**

## Auto-discovery routes and inter-PE tunnels

- **Auto-discovery routes received by a PE provides the PE connected to a MVPN site with:**
  - (a) the information about the identity of other PEs connected to the sites of that MVPN
    - Auto-discovery route originated by a PE carries the identity of that PE (IP address of the PE)
  - (b) the identity of the tunnels used by other PEs for sending multicast traffic of that MVPN
    - Auto-discovery route originated by a PE for a particular MVPN carries the identity of the inter-PE tunnel that the PE will use to send to other PEs that have sites of that MVPN multicast traffic coming from the sources in the MVPN site(s) connected to that PE
    - Tunnel identity is carried in the BGP PMSI Tunnel attribute of the auto-discovery route
    - Tunnel identity includes the type of the Tunnel signaling protocol (e.g., RSVP-TE, LDP, etc...)
- **Combination of (a) and (b) provides sufficient information for tunnel signaling**

## Auto-discovery routes and inter-PE tunnels: example



## Binding multicast streams to inter-PE tunnels: Default vs Selective tunnels

- **Default tunnel advertised by a given PE carries MVPN multicast streams:**
  - from all the sources in the MVPN site(s) connected to the PE
  - all the multicast streams originated by the sources
  - to all the PEs connected to all other sites of that MVPN
    - even if some of these sites have no actual receivers for the multicast streams, HOWEVER
    - CEs in the sites with no actual receivers do not receive the multicast streams
- **Selective tunnel advertised by a given PE carries MVPN multicast streams:**
  - only from a particular source(s) in the MVPN site(s) connected to the PE
  - only a subset of multicast streams originated by the sources
  - to only the PEs connected to the other sites of that MVPN that have actual receivers for the multicast streams

**Default tunnels require less forwarding state than Selective tunnels.**

**Selective tunnels are more bandwidth efficient than Default tunnels.**

## More on Default and Selective tunnels

- **Both Default and Selective tunnels are established using the information carried in BGP auto-discovery routes**
- **BGP auto-discovery routes for Selective tunnels carry additional information about specific MVPN multicast sources (S) and groups (G) that would be carried over the Selective tunnels**
  - Binds a particular <S,G> of a particular MVPN to a particular Selective tunnel
- **Creating selective tunnel for (S,G) of a given MVPN is controlled solely by the PE connected to the site that contains S**

## More on inter-PE tunnels: tunneling technologies

- **Available tunneling technologies:**
  - MPLS-based: P2MP LSP with RSVP-TE, P2MP LSP with LDP, Ingress replication
  - GRE-based: PIM-SSM with GRE encapsulation, PIM-SM with GRE encapsulation, PIM-Bidir with GRE encapsulation
- **Different MVPNs within the same Service Provider may use different tunneling technologies**
- **For a given inter-AS/inter-provider MVPN (an MVPN that spans multiple ASes/providers) each provider may use different tunneling technology**

## More on inter-PE tunnels: aggregation

- **Aggregate multiple P2MP LSPs using P2MP LSP hierarchy**
  - Multiple P2MP LSPs (rooted at the same PE) may be “nested” inside a single (outer) P2MP LSP
  - Applicable to P2MP LSPs used for both Default and Selective tunnels
  - Not constrained by MVPN boundaries
    - P2MP LSPs of multiple MVPNs could be aggregated into a single P2MP LSP
  - Could use aggregation with partial congruency
    - If the aggregated P2MP LSPs are only partially congruent with each other
- **Aggregate multiple (S,G) of a given MVPN into a single Selective tunnel**
  - As long as all the sources are in the sites connected to the PE that creates the Selective tunnel
- **Results in improved scalability by reducing both forwarding plane and control plane overhead**

## Agenda

- BGP/MPLS MVPN – what are the goals ?
- Supporting PIM-SM in SSM mode MVPNs
- ➔ ▪ Supporting PIM-SM in ASM mode MVPNs
- Summary



## Supporting PIM-SM in ASM mode MVPNs

- **Supporting PIM-SM in ASM mode MVPNs requires:**
  - A mechanism to carry MVPN multicast routing information from the PEs connected to the sites that contain the Receivers to the PE connected to the sites that contain the Sources
    - Build upon the mechanism used for PIM-SM in SSM mode MVPNs
    - Described in the following slides
  - A mechanism to carry multicast traffic from the PE connected to the sites that contains the Sources to the PEs connected to the sites that contain the Receivers
    - The same as the one used for PIM-SM in SSM mode MVPNs

**Observation: the same decomposition as for supporting PIM-SM in SSM mode MVPNs**

## Supporting PIM-SM in the ASM mode MVPNs: few observations

### In the context of plain IP multicast:

- **The sole purpose of joining RP Tree (RPT) is for the receivers to discover the sources**
  - As RP knows about all active sources
    - Designated Routers connected to active sources register the active sources with RP (using PIM Register)
- **Switching from RP Tree (RPT) to Shortest Path Tree (SPT) usually occurs as soon as a receiver discovers a source**
  - Usually on the first packet received from the source
- **RPT/SPT interaction introduces a fair amount of additional complexity**
  - PIM-SM in the SSM mode is (significantly) simpler than PIM-SM in the ASM mode

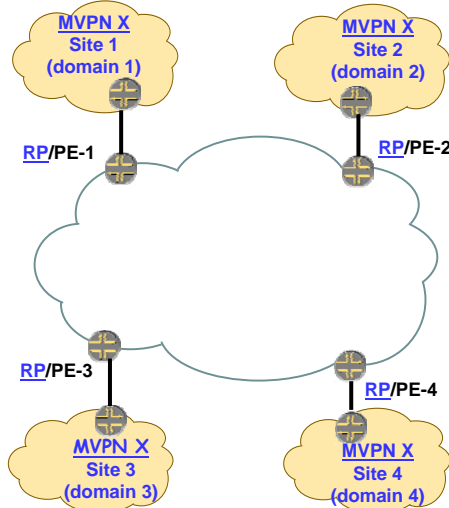
## Supporting PIM-SM in ASM mode MVPNs: key ideas

- **Within each MVPN use the existing IP multicast mechanisms to enable RP(s) of that MVPN to discover active sources (S,G) of that MVPN**
- **Bring the information about active sources (S,G) to one or more PEs by either:**
  - Option 1:** Co-locating MVPN RP(s) with PEs
  - OR
  - Option 2:** Using the existing IP multicast mechanisms to communicate information about active sources (S,G) from RP(s) to one or more PEs
- **Use BGP Source Active auto-discovery routes to distribute information about active sources (S,G) within a given MVPN among all the PEs connected to the sites of that MVPN**
- **Use BGP C-multicast routes to inform PEs connected to active sources that there are receivers connected to some other PEs**
  - Just like with PIM-SM in SSM mode

## BGP Source Active auto-discovery routes

- **Source Active auto-discovery routes are carried in Multiprotocol BGP (RFC4760) using MCAST-VPN NLRI**
  - The same NLRI as used by C-multicast routes and auto-discovery routes
- **Source Active auto-discovery route NLRI contains:**
  - Multicast Source (S), Multicast Group (G)
  - Route Distinguisher (RD)
    - Needed to support MVPNs that may use the same address space (just like with unicast)
- **Route Targets used to constrain distribution of Source Active auto-discovery routes of a given MVPN can be the same as the MVPN import/export Route Target**
  - MVPN import/export Route Targets can be the same as Route Targets used by unicast 2547 VPNs
- **Re-use the existing BGP mechanisms (e.g., extended communities, Route Target constraint, Route Reflectors, etc...)**

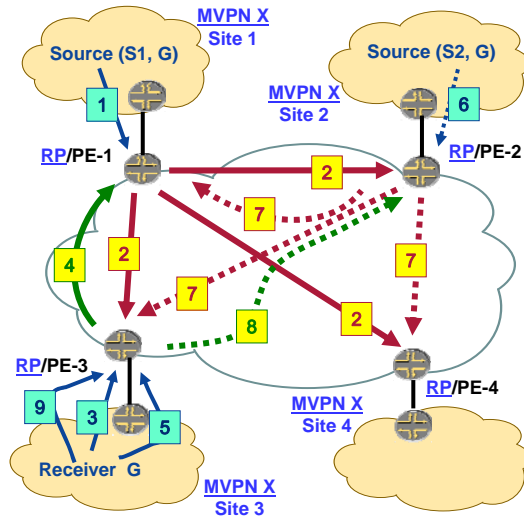
## Option 1: Colocated RP/PE solution – MVPN as a collection of (interconnected) multicast domains



- Each (multicast) domain consists of all the sites of a given MVPN connected to a given PE
- MVPN = set of multicast domains interconnected by the MVPN Service Provider(s) infrastructure
- A given PE acts as an (MVPN) RP for all the sites of a given MVPN connected to that PE – *colocated RP/PE*
  - Distinct RP instance per directly connected MVPN on the PE
- Plain PIM-SM in the ASM mode procedures for all the sites of a given MVPN connected to a given PE, including CE-PE interaction
  - Plain PIM-SM in the ASM mode within each (multicast) domain
- But what about inter-site (inter-domain) procedures among the MVPN sites connected to different PEs ?
  - See the following slide...

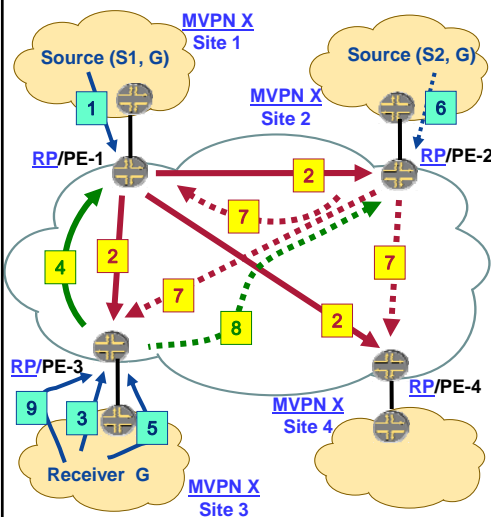


### Option 1: example



1. PIM Register (S1, G)
2. Source Active auto-discovery (S1, G)
3. PIM Join (\*, G)
4. C-Multicast (S1, G)
5. PIM Join (S1, G)
6. PIM Register (S2, G)
7. Source Active auto-discovery (S2, G)
8. C-Multicast (S2, G)
9. PIM Join (S2, G)

### Option 1: Summary



- Each RP/PE of a given MVPN discovers active sources (S,G) within the MVPN sites (directly connected to the RP/PE by using plain PIM)
  - (1) PIM Register (S1, G), (6) PIM Register (S2, G)
- RP/PE connected to the site that contains an active source (S,G) informs all other PEs that have sites of that MVPN about the active source (S,G) by originating Source Active auto-discovery route for the active source
  - (2) Source Active (S1, G), (7) Source Active (S2, G)
- Receivers within each MVPN site use plain PIM to inform the PE connected to the site that they want to receive traffic for a particular group G
  - (3) PIM Join (\*, G)
- When a PE receives Join(\*,G) from one of its directly connected CEs, the PE converts it into C-multicast routes, one per each received Source Active auto-discovery route that has G
  - (4) C-Multicast route (S1, G), (8) C-Multicast route (S2, G)
  - Informs the PE connected to the active source (S,G) that there are receivers for (S,G) connected to some other PEs
- When a receiver switches from Shared Tree (RPT) to Source Tree (SPT), this switch is localized to the site that contains the receiver and the PE connected to the site
  - (5) PIM Join (S1, G), (9) PIM Join (S2, G)

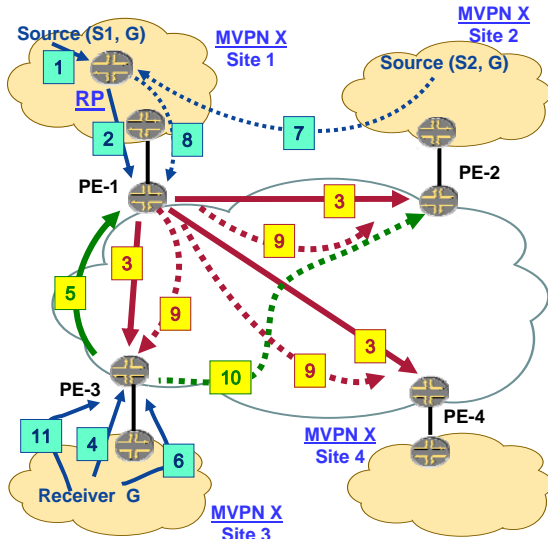
## Option 1: Summary (cont.)

- **Requires every PE that has one or more sites of a given MVPN connected to it to act as an RP for that MVPN**
- **Suitable if an MVPN customer wants to outsource its RP infrastructure to the service provider**
- **Problematic if an MVPN customer wants to retain its own RP infrastructure**
- **See next slides for other options...**

## Option 2a/2b

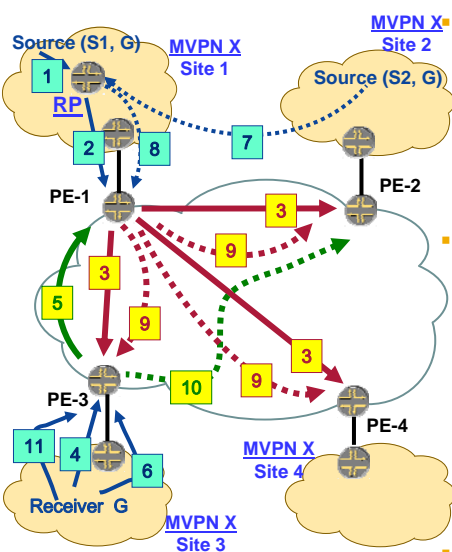
- **MVPN customer maintains its own RP infrastructure**
- **Use the existing IP multicast mechanisms to communicate information about active sources (S,G) from MVPN RP(s) to one or more PEs**
  - These PEs maintain information about active sources (S,G) for a given MVPN
- **Option 2a: use MSDP between MVPN RPs and PEs**
- **Option 2b: use PIM Register between MVPN RPs and PEs**
- **PEs do NOT act as MVPN RPs**
  - PEs do not receive PIM Register from any of the MVPN Designated Routers (DRs)
  - With Option 2a MSDP SAs flow only from RP to PE, but not from PE to RP
  - With Option 2b PIM Registers flow only from RP to PE, but not from PE to RP

## Option 2a: example (MSDP between RP and PE)



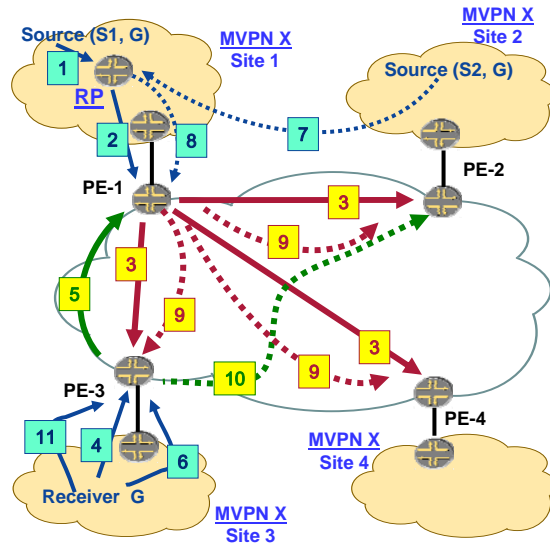
1. PIM Register (S1, G)
2. MSDP SA (S1, G)
3. Source Active auto-discovery (S1, G)
4. PIM Join (\*, G)
5. C-Multicast (S1, G)
6. PIM Join (S1, G)
7. PIM Register (S2, G)
8. MSDP SA (S2, G)
9. Source Active auto-discovery (S2, G)
10. C-Multicast (S2, G)
11. PIM Join (S2, G)

## Option 2a: Summary (MSDP between RP and PE)



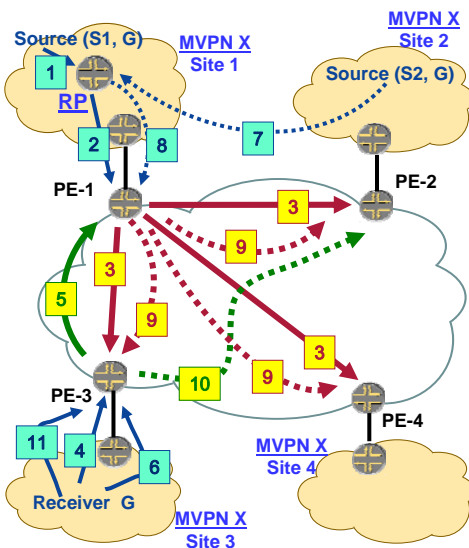
- RPs within each MVPN discover information about MVPN active sources (S,G) by using plain IP multicast mechanism (PIM Register messages)
  - Does not require any multicast within the service provider
  - (1) PIM Register (S1, G), (7) PIM Register (S2, G)
- PE obtains from an RP of a given MVPN the information about active sources (S,G) within that MVPN using plain IP multicast mechanisms - MSDP between PE and RP
  - Distinct MSDP instance on the PE per each distinct MVPN
  - MSDP Source Active (SA) advertisements flow from RP to PE, but NOT from PE to RP
  - (2) MSDP SA (S1, G), (8) MSDP SA (S2, G)
  - PEs do NOT maintain MSDP peering with each other
- The rest is the same as with Option 1

## Option 2b: example (PIM Register between RP and PE)



1. PIM Register (S1, G)
2. PIM Register (S1, G)
3. Source Active auto-discovery (S1, G)
4. PIM Join (\*, G)
5. C-Multicast (S1, G)
6. PIM Join (S1, G)
7. PIM Register (S2, G)
8. PIM Register (S2, G)
9. Source Active auto-discovery (S2, G)
10. C-Multicast (S2, G)
11. PIM Join (S2, G)

## Option 2b: Summary (PIM Register between RP and PE)



- RPs within each MVPN discover information about MVPN active sources (S,G) by using plain IP multicast mechanisms (PIM Register messages)
  - Does not require any multicast within the service provider
  - (1) PIM Register (S1, G), (7) PIM Register (S2, G)
- PE obtains from an RP of a given MVPN the information about active sources (S,G) within that MVPN using plain IP multicast mechanisms – PIM Register between PE and RP
  - PIM Register messages flow from RP to PE, but NOT from PE to RP
  - (2) PIM Register (S1, G), (8) PIM Register (S2, G)
- The rest is the same as with Option 1

## Option 2a/2b: Summary

- **Works well if an MVPN customer wants to have full control over its own RP infrastructure**
  - Supports Anycast RP in customer's RP infrastructure
  - Supports BSR, Auto-RP in customer's RP infrastructure
- **Does this option make sense if a customer wants to completely outsource the RP infrastructure ?**
  - NO, as this option assumes that none of the PEs act as an RP

## Options 1, 2a/2b comparison

- **Among MVPN sites the (multicast) traffic is ALWAYS carried over Shortest Path (SPT) trees**
  - Inter-site (multicast) traffic never flows through customer's RP
- **All these options have exactly the same procedures for:**
  - Originating and receiving BGP Source Active auto-discovery routes for active sources (S,G)
  - Originating and receiving BGP C-multicast routes that carry (S,G)
  - Handling PIM messages received by PEs from the directly connected CEs
- **The main difference is whether a PE acts as a fully functional MVPN RP**
  - Yes with Option 1
  - No with Option 2a/2b

## Agenda

- BGP/MPLS MVPN – what are the goals ?
- Supporting PIM-SM in SSM mode MVPNs
- Supporting PIM-SM in ASM mode MVPNs
- ➔ ▪ Summary



## BGP/MPLS MVPN – Summary

- **Extends 2547 VPN service offering to include support for IP multicast for 2547 VPN customers**
- **Follows the same architecture/model as 2547 VPN unicast**
  - Uniform control plane to support both unicast and multicast
  - Eliminates the need to have the Virtual Router (VR) model for multicast and the 2547 model for unicast
- **Re-uses 2547 VPN unicast mechanisms: BGP, MPLS**
  - With extensions, as necessary
  - Common set of mechanisms to support both unicast and multicast
- **Retains as much as possible the flexibility and scalability of 2547 VPN unicast**

## Suggested reading

- draft-ietf-l3vpn-ppvnp-mcast-reqts
- draft-ietf-l3vpn-2547bis-mcast
- draft-ietf-l3vpn-2547bis-mcast-bgp
- draft-morin-l3vpn-mvpn-considerations

Juniper *your* Net™