

Multicast VPN: Multicast Distribution Trees Subaddress Family Identifier Information

Last updated: September 2007

Multicast VPNs provide a method to separate private multicast traffic over a shared core network. This document covers interoperability and deployment aspects of the creation of Multicast VPN (MVPN) Multicast Distribution Trees (MDTs) over the shared core with regards to the introduction of MDT Subaddress Family Identifier Information (SAFI) in Cisco IOS Software Release 12.0(29)S. A basic understanding of MVPN is assumed.

The Need for MDT SAFI

In a single Autonomous System (AS) environment, if the default MVPN-MDT is using Protocol Independent Multicast, Sparse mode (PIM-SM) with Rendezvous Points (RPs), then PIM over the Multicast Tunnel Interface (MTI) will establish adjacencies because the source PE and receiver PE discover each other via the RP. In this case, the local PE (source PE) sends Register messages to the RP, which then builds a Shortest-Path Tree (SPT) towards the source PE. The remote PE which acts as a receiver for the MDT multicast group, sends*, G joins toward the RP and joins the distribution tree for that group.

However, if the default MDT group is PIM- Source Specific Multicast (PIM-SSM) rather than PIM-SM as described above, the receiver PE needs information about the source PE and the default MDT group. This information is used to send S, G joins towards the source PE, building a distribution tree from the source PE without the need for a RP. The IP address of the source PE and default MDT group is sent via Border Gateway Protocol (BGP).

There are two ways in which BGP can send this information:

Cisco IOS Software Version	BGP Advertisement Method
Pre-Release 12.0(29)S Pre-Release 12.2(31)SB	Extended communities and Type 2 RD
Release 12.0(29)S and above Release 12.2(31)SB and above Release 12.2(33)SRA1 and above IOS XR 3.5	BGP address family MDT SAFI

BGP Extended Community was used as an interim solution before IETF submission. It has certain limitations, it cannot be used in Inter-AS scenarios (as the attribute is BGP non-transitive) and that it is a proprietary Cisco solution.

The new address family based advertisement method, documented in draft-nalawade-idr-mdt-safi overcomes these limitations.

Pre 29S: BGP Extended Community

The PE loopback (source address) information is sent as a VPNv4 prefix using RD (Route Distinguisher) Type 2 (to distinguish it from unicast VPNv4 prefixes). The MDT group address expressed in a BGP Extended Community. Using the combination of the embedded source in the VPNv4 address and the multicast group in the extended community, PE routers in the same Multicast Virtual Routing and Forwarding (MVRF) can setup SSM trees to each other.

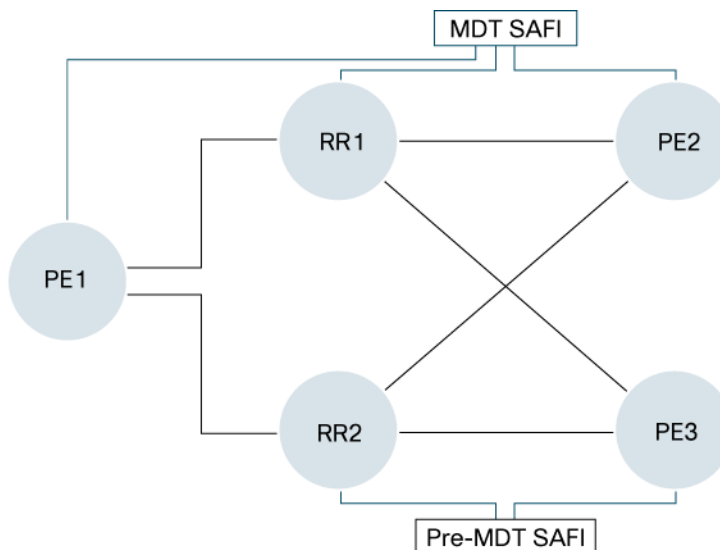
There are two basic issues with this approach, use of RD Type 2 is not standard, and the BGP attribute is non-transitive, and cannot work for Inter-AS MVPN. In “BGP/MPLS IP VPNs” (RFC4364), RD Type 2 is designated for the representation of 4 byte AS numbers (RFC4893). This overlap in usage can cause collisions between the RD Type 2 MVPN MDT advertisement and a unicast VPNv4 advertisement using RD Type 2 for encoding the 4 byte AS number.

Post 29S: MDT SAFI

The MDT address family was introduced in Cisco IOS Software Release 12.0(29)S and IOS-XR Release 3.5. The same information, the source PE and MDT group address, is passed to PIM now using the BGP MDT-SAFI update. The RD type and format is now the same as the VRF's unicast VPNv4 advertisements. BGP will do a best path on the MDT updates before passing the information to PIM.

To prevent backwards compatibility problems, Cisco IOS Software BGP has introduced functionality to allow communicating the old style updates with peers that are unable to understand the new SAFI address family.

Consider the following diagram:



PE1, RR1 and PE2 are running the MDT-SAFI code, RR2 and PE3 are running old code which is unaware of MDT SAFI.

PE1 negotiates the new MDT SAFI address family with RR1, and will do backwards compatibility with RR2. PE1 will create a new style BGP update to RR1 and an old style to RR2.

PE2 will get two different style updates between RR1 and RR2. The old style update which is received from RR2 is converted to MDT SAFI by BGP on PE2. The BGP best path is run on both

old (now converted) and new style updates. The result of BGP best path calculation is then forwarded to PIM.

It should be noted that PE3 will only receive old style MDT updates from RR2 and RR1. RR1 will convert down to old style MDT updates before advertising to PE3.

For IOS-XR, the MDT-SAFI is the only supported format and backward compatibility translation is not an option.

While backwards compatibility can help in deployment, if the MDT SAFI is being used then it is better to convert the network fully to use MDT SAFI. The reason for this is to reduce the additional conversion load on the route-reflector and PE routers.

As the new MDT-SAFI does not use BGP route-target extended communities, the regular extended community methods to filter these updates no longer apply. A new keyword under route-maps ('match mdt-group <acl>') has been added to filter on the MDT group address by access-lists. This route-map can be applied inbound or outbound to the IPv4 MDT address-family neighbor configuration.

Show Commands

BGP will automatically convert old style updates received to the new format. For example, the advertisement from 12.0.0.9 was originally received as an old style update but was automatically converted to new style.

```
RR1#show ip bgp ipv4 mdt all
BGP table version is 12, local router ID is 12.0.0.3
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
                r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop          Metric LocPrf Weight Path
Route Distinguisher: 13979:2
*>i12.0.0.2/32      12.0.0.2          0      100      0 ?
* I                12.0.0.2          0      100      0 ?
Route Distinguisher: 13979:4
* i12.0.0.4/32     12.0.0.4          0      100      0 ?
*>i                12.0.0.4          0      100      0 ?
Route Distinguisher: 13979:9
*>i12.0.0.9/32     12.0.0.9          0      100      0 ?
* i                12.0.0.9          0      100      0 ?
```

The following command shows the PIM MDT table that is learned from BGP. In the below output we are seeing new MDT SAFI updates.

```
RR1#show ip pim mdt bgp
MDT (Route Distinguisher + IPv4)          Router ID      Next Hop
MDT group 232.0.3.32
  13979:2:12.0.0.2                        12.0.0.2      12.0.0.2
  13979:4:12.0.0.4                        12.0.0.4      12.0.0.4
  13979:9:12.0.0.9                        12.0.0.9      12.0.0.9
```

On RR2, which is running old code, the old style MDT updates are received by BGP and directly forwarded to the PIM sub-system.

```
RR2#show ip bgp vpnv4 all | be Route Distinguisher: 2:
Route Distinguisher: 2:13979:2
*>i12.0.0.2/32  12.0.0.2      0    100    0 ?
* I           12.0.0.2      0 100    0 ?
Route Distinguisher: 2:13979:4
* i12.0.0.4/32  12.0.0.4      0    100    0 ?
*>i           12.0.0.4      0 100    0 ?
Route Distinguisher: 2:13979:9
* i12.0.0.9/32  12.0.0.9      0    100    0 ?
*>i           12.0.0.9      0    100    0 ?
```

```
RR2#show ip pim mdt bgp
Peer (Route Distinguisher + IPv4) Next Hop
MDT group 232.0.3.32
  2:13979:2:12.0.0.2                12.0.0.2
  2:13979:4:12.0.0.4                12.0.0.4
  2:13979:9:12.0.0.9                12.0.0.9
```

Configuration

In pre-MDT SAFI code, apart from the BGP configuration required for supporting MPLS VPNS, additional configuration for MVPN was not required. As part of the new MDT SAFI code the BGP MDT SAFI address family now needs to be explicitly configured for MVPN BGP neighbors. It should be noted that neighbors that do not support the new MDT SAFI would still need it enabled to support backward translation.

Pre-MDT SAFI configuration:

```
ip vrf <vrfName>
  rd <rd>
  route-target <rt>
  mdt default <mdt group>
!
ip multicast-routing
ip multicast-routing vrf <vrfName>
!
router bgp <AS>
...
  address-family vpnv4
    neighbor <iBGP peer> activate
    neighbor <iBGP peer> send-community-both
```

New MDT SAFI configuration:

```
ip vrf <vrfName>
  rd <rd>
  route-target <rt>
  mdt default <mdt group>
!
ip multicast-routing
ip multicast-routing vrf <vrfName>
!
router bgp <AS>
...
  address-family vpnv4
    neighbor <iBGP peer> activate
    neighbor <iBGP peer> send-community-both
!
address-family ipv4 mdt
  neighbor <iBGP peer> activate
  neighbor <iBGP peer> send-community-both
```

Auto Migration to MDT SAFI

For existing VPNv4 neighbors, auto configuration of MDT SAFI neighbors, now occurs upon bootup based on the presence of existing MDT configuration ('mdt default x.x.x.x' under a VRF). Auto-migration was introduced in Releases 12.0(30)S3, 12.0(31)S1, and 12.0(31.1)S. Pre-SAFI configurations will be automatically converted to MDT SAFI configuration upon bootup to ease migration. When manually configuring a VPNv4 neighbor in BGP and a default MDT is configured, a similar neighbor in the ipv4 mdt address-family will also be automatically configured.

Also note that because there is no VRF configuration on route-reflectors the MDT SAFI auto-migration will not be triggered and will have to be manually entered.

Recommendations

Configure MDT SAFI

Upon reload, MDT SAFI will be configured automatically on PE devices, it is recommended to configure MDT SAFI on the route-reflectors (as the migration will not be automatic). Having a uniform MDT transmission method will reduce processing time on the routers as conversion will not need to be done.

Even though the benefits of MDT SAFI are for SSM tree building, MDT SAFI must also be configured when using MVPN with the default MDT group in PIM sparse-mode. From the multicast point of view, the new BGP AF does not need to be configured for MVPN to work with a PIM-SM core. However, the fix for CSCef97738, which was introduced in Release 12.0(30)S1, mandates that the new address family be configured in order to create the MTI interface in certain conditions. Without this notification, the multicast tunnel would not be created, and hence MVPN will not work (even with PIM-SM). It should be noted, that in the case of PIM-SM in the core, while the MDT SAFI needs to be configured, it is not required to actually receive MDT advertisements to establish MDT PIM relationships.

BGP Peer Extended Community Sending

For backward compatible sessions, extended community sending for the MDT SAFI peers must be enabled. In a pure BGP SAFI environment there is no need to configure extended communities explicitly for MVPN. However, extended communities will definitely be needed for VPNv4 iBGP sessions to relay the route-target. To recap, in a hybrid (SAFI and non-SAFI) environment, it is necessary to configure extended communities sending for MDT SAFI neighbors.

Network Upgrade and Conversion

When moving from non-SAFI code to SAFI capable code, care has to be taken to make as little impact as possible to the MVPN service. The unicast service will not be affected, other than the outage due to the reload and recovery.

Upgrade MVPN PE's to MDT SAFI capable code. The configs will be auto-migrated.

After the PEs have been upgraded, upgrade the RR and enable MDT SAFI for all peers providing MVPN service. Note that enabling/disabling MDT SAFI will reset the BGP peer relationship for all address families, and thus loss of routing information may occur. In the case of a multi-homed route-reflector scenario:

One of the RR's has to be upgraded and configured last. The upgraded PEs will use this RR to relay MDT advertisements while the other RRs are being upgraded.

References

CSCef97738

BGP/MVPN: incorrect update-source is passed from BGP to MVPN

CSCei25442

BGP MVPN - sanitize notification to PIM database

CSCei25454

BGP: The BGP IPv4 MDT table may receive RD type 2 prefixes

[CSCeg52240](#)

BGP: Installing old MDT extcomm as MDT prefix in MDT SAFI table

[CSCsf31082](#)

MVPN: After clearing iBGP, default MDT SSM loses other PEs as PIM neighbors

[CSCee40925](#)

Auto-configure MDT SAFI from vpnv4 configs for MVPN backward compatibility.

[CSCek38025](#)

MDT interoperability with old image is not working

[CSCsc35663](#)

MVPN: mal-formatted bgp updates with mdt AF

[[MDT SAFI](#)] G. Nalawade, et. Al., 'MDT SAFI', draft-nalawade-idr-mdt-safi

[[2547bis-mcast](#)] E. Rosen, R. Aggarwal, et. Al., 'Multicast in MPLS/BGP IP VPNs', draft-ietf-l3vpn-2547bis-mcast

Multicast VPN Data Sheet

http://www.cisco.com/en/US/technologies/tk648/tk828/tk363/technologies_white_paper090aecd802aea84.shtml



Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NLTS (6387)
Fax: 408 527-0689

Asia Pacific Headquarters
Cisco Systems, Inc.
155 Robinson Road
#29-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Europe Headquarters
Cisco Systems International BV
Hearstorgpark
Hearstorgweg 13-19
1101 CH Amsterdam
The Netherlands
www.europe.cisco.com
Tel: +31 20 60 020 0/91
Fax: +31 20 657 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

©2007 Cisco Systems, Inc. All rights reserved. CCVP, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Register, Airnet, BPX, Catalyst, CCNA, CCDP, CCIE, CCR, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast, Step, Follow Me Browsing, FormShare, Go to Drive, HomeLink, Internet Quotient, IOS, IPPhone, IPTV, IQ Expertise, the IQ logo, IQ Not Ready, Note Scorecard, QuickStudy, iQoSStream, iInlays, Meeting Place, MGX, Networking Academy, Network Registrar, Packet, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (9705R)