

MDT SAFI  
[draft-nalawade-idr-mdt-safi-03.txt](#)

## **1. Status of this Memo**

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

## **2. Copyright Notice**

Copyright (C) The Internet Society (2005). All Rights Reserved.

## **3. Abstract**

There is a need for transporting Multicast Tunnel attributes within and across Autonomous Systems. This draft defines a new Address-Family that can be used to do the distribution.

## **4. Introduction**

Two end-points of a Multicast Tunnel need to know the end-point information and its binding to a network address and the Multicast Tunnel used for the respective VRF at the remote PE. Normally, Tunnel endpoint addresses can be statically configured. But in case of a large network with large number of VPNs where there may be a need for a large number of endpoints and a large number of VRFs, the amount of information that needs to be exchanged and maintained between PEs for Multicast Tunnels for VPNs, is quite large. It therefore needs a mechanism that can maintain and support Multicast Tunnel based VPNs in a flexible, scalable manner. Also, in inter-AS and inter-provider scenarios, this mechanism needs to be supported across autonomous systems and provider domains.

## 5. The MDT SAFI

A new Subsequent-Address Family called the MDT SAFI is being defined. The NLRI for this SAFI, will contain the address of the nexthop which will be used by the multicast source PE to send the PIM Join for the default MDT address to.

The NLRI format is 8-byte-RD:IPv4-address followed by the MDT group address. i.e. The MP\_REACH attribute for this SAFI will contain one or more tuples of the following form :

```
+-----+
|                                     |
|  RD:IPv4-address (12 octets)      |
|                                     |
+-----+
|  MDT Group-address (4 octets)    |
+-----+
```

where :

Route-Distinguisher : is the RD of the VRF to which this MDT attribute belongs.

MDT Group Address : is the Group-address of the MDT-Group that a VRF is associated to. This can be variable length. But for

IPV4 addresses - this will be 4 octets.



## 6. The Connector Attribute

An Optional Transitive Connector attribute [[BGP-CONN](#)] will be used to transport the address of the originating PE router unchanged to the remote PE router which is participating in the same MVPN.

The format of this attribute is a 2-octet Type field followed by the Value. Type 1 of this attribute defines that this is an IPv4 unicast nexthop set by the PE which advertises the CE-learnt prefixes to its peers.

The attribute contains one or more tuples of the form :

```
+-----+
|                                     |
|  Type (2 octets)                   |
+-----+
|                                     |
|  Value (Variable)                  |
|                                     |
+-----+
```

where :

Type : is the Type of the data contained in this TLV.

Value : is a variable length field as defined by the Type.

When the Type is 1, the value will contain the IPv4 address of the PE sourcing the CE-learnt VPNv4 prefixes. This would be the same address this PE uses to set itself as the nexthop. This attribute will accompany the VPNv4 prefix advertisement.

## 7. Capability Advertisement

A BGP speaker that wishes to exchange the MDT SAFI, MUST use the MP\_EXT Capability Code as defined in [[BGP-MP](#)], to advertise the corresponding (AFI, SAFI) pair.

A BGP speaker MAY participate in the distribution of MDT information.



## 8. Operation

A BGP Speaker that receives the Capability for the MDT SAFI, MAY advertise the MDT SAFI prefixes to that peer. The prefixes in the MDT SAFI are populated by the PEs that advertise their CE-learned prefixes.

## 9. Applicability Statement

Multicast Tunnels are built between Provider Edge (PE) routers to allow multicast communication between different sites of a VPN. The MT tunnel has a destination MDT group address that is unique to the VPN. All routers that act as PE's and are configured for a specific VPN join the VPN MDT multicast group in the backbone of the provider network to be able to receive each other's packets. Each router is also a sender to the MDT group. In PIM SSM mode, the following procedure is followed.

A Multicast tunnel is setup between the PEs in one or more VPN-Providers networks. Over the Multicast tunnel we create PIM neighbor's. The IP address of the PIM neighbor that is seen over the Multicast tunnel depends on the configured address of the Tunnel endpoint. This can either be an unnumbered address from a different interface or a configured address on the Tunnel itself. The PE router that does an RPF check on a VPN source can find which Tunnel the source is on, but may not know what PIM neighbor to target on that tunnel. Therefore we need a way to connect the BGP VPNv4 prefix to the PIM neighbor on the tunnel to allow the RPF check to succeed.

Suppose we want to join to a source that is behind another VPN site. We do an RPF lookup on the source address in the VPNv4 unicast table on this PE. The RPF lookup will return a connected next-hop and interface to use to reach the source. The returned next-hop may not be the neighbor on the Multicast tunnel. This can be due to the next-hop being rewritten by BGP Route Reflectors (RR) or crossing AS's. Therefore we don't know which PIM neighbor to target as upstream neighbor in the PIM join.

defines a new attribute called the BGP Connector attribute. This We propose sending the Originating PE's IP address through the BGP Connector Attribute. It will be sent as the value field when the BGP Connector attribute contains Type 1. This is the Multicast Tunnel's IP address which is used to establish the PIM neighbor relationship on the Multicast tunnel. This attribute is attached to all the BGP VPNv4 prefixes used for reaching to multicast sources in the Customer's network. The PE router that was able to successfully RPF on a BGP VPNv4 prefix will use the IP address learned from the

connected attribute to find the PIM neighbor on the Multicast tunnel.

[draft-nalawade-idr-mdt-safi-03.txt](#)

4]

[Page

## **10. Security Considerations**

This extension to BGP does not change the underlying security issues.

## **11. Acknowledgements**

The authors would like to thank IJsbrand Wijnands for his contribution towards the Connector Attribute needed for the IPv4-MDT SAFI. The authors would also like to thank Shyam Suri, Ruchi Kapoor, Shyam Suri, Yiqun Cai, Dan Tappan, Jennifer Li, Yi Chou, Arjen Boers for their comments and contributions. We would also like to thank Bhavani Parise for his review and comments.

## **12. Normative References**

[BGP-4] Rekhter, Y. and T. Li (editors), "A Border Gateway Protocol 4 (BGP-4)", Internet Draft [draft-ietf-idr-bgp4-26.txt](#), April 2005.

[BGP-CAP] Chandra, R., Scudder, J., "Capabilities Advertisement with BGP-4", [draft-ietf-idr-rfc2842bis-02.txt](#), April 2002.

[BGP-SSA] Kapoor R., Nalawade G., BGPv4 SAFI-Specific Attribute , [draft-nalawade-kapoor-bgp-ssa-00.txt](#), January 2006.

[MULTI-BGP] Bates et al, Multiprotocol Extensions for BGP-4, [draft-ietf-idr-rfc2858bis-02.txt](#), work in progress.

[BGP-CONN] Nalawade Gargi, Kapoor Ruchi "BGPv4 Connector Attribute", OCT-05, <[draft-nalawade-bgp-connector-00.txt](#)>, work in progress.

## **13. Author's Addresses**

Gargi Nalawade  
170 Tasman Drive  
San Jose, CA, 95134  
E-mail: [gargi@cisco.com](mailto:gargi@cisco.com)

Arjun Sreekantiah  
170 Tasman Drive  
San Jose, CA, 95134  
E-mail: [asreekan@cisco.com](mailto:asreekan@cisco.com)

## **14. Intellectual Property Statement**

Cisco Systems may seek patent or other intellectual property protection for some of all of the technologies disclosed in this document. If any standards arising from this document are or become protected by one or more patents assigned to Cisco Systems, Cisco





intends to disclose those patents and license them on reasonable and non-discriminatory terms.

#### **15. Full Copyright Statement**

Copyright (C) The Internet Society (2005). All Rights Reserved.

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### **16. Expiration Date**

This memo is filed as [<draft-nalawade-idr-mdt-safi-03.txt>](#), and expires April, 2006.

