Internet Engineering Task Force (IETF)
Request for Comments: 6420
Category: Standards Track
ISSN: 2070-1721

Y. Cai
H. Ou
Cisco Systems, Inc.

November 2011

PIM Multi-Topology ID (MT-ID) Join Attribute

Abstract

This document introduces a new type of PIM Join Attribute that extends PIM signaling to identify a topology that should be used when constructing a particular multicast distribution tree.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6420.

Table of Contents

1. Introduction

Some unicast protocols, such as OSPF and IS-IS, allow a single network to be viewed as multiple topologies [RFC4915] [RFC5120]. Deploying multi-topology (MT) routing allows different paths through the network to be selected to support different traffic or to offer protection paths in the event of failures.

PIM [RFC4601] employs a technique known as Reverse Path Forwarding (RPF) to construct forwarding trees between multicast sources and receivers. The procedure of RPF uses topology information provided by routing protocols, such as OSPF and IS-IS. Using the PIM MT-ID Join Attribute specified in this document enables PIM to access the multiple topologies created by the routing protocols and construct multicast forwarding trees using separate network paths even when the roots of the trees are the same.

This capability would allow for an improvement to the resilience of multicast applications. For instance, a multicast stream can be duplicated and transported using two source trees, (S1, G1) and (S1, G2), simultaneously. By using MT-capable unicast routing protocols and procedures described in this document, it is possible to construct two source trees for (S1, G1) and (S1, G2) in such a way that they do not share any transit network segment. As a result, a single network failure will not cause any loss to the stream.

This document introduces a new type of PIM Join Attribute [RFC5384], named "MT-ID Join Attribute". It is used to encode the numerical identity of the topology PIM uses when performing RPF for the forwarding tree that is being joined. This document also specifies procedures and rules to process the attribute and resolve conflicts arising from mismatches in capabilities to support the attribute or the value of the attribute.

This document does not introduce any change to the RPF check procedure used to verify the incoming interface when a packet is forwarded as defined in [RFC4601]. For example, to use the capability described by this document, an application can choose to use group addresses, and/or source addresses, to identify a unique multicast stream. It might further need to perform the functions of splitting and merging. However, the detailed processing is beyond the scope of the document.

In the rest of the document, the MT-ID Join Attribute will be referred to as "MT-ID".

2. Terminology

The following acronyms are frequently used in the document.

- RPF: RPF stands for "Reverse Path Forwarding". A PIM router performs RPF for two purposes. When building a forwarding tree, a PIM router identifies an interface (the RPF interface) and an upstream PIM neighbor (the RPF neighbor) to which to send PIM Joins. Upon receiving a data packet, a PIM router verifies if the packet arrives from the expected incoming interface (aka RPF check) before deciding whether or not to replicate the packets.

- RPF Topology: An RPF topology is a collection of routes that a PIM router uses for RPF. One or more RPF topologies may be created on a PIM router.

- MT: MT stands for "Multi-Topology" in this document. Sometimes it is also referred to as "multi-topology routing". In the context of PIM, MT refers to the capability of building and maintaining multiple RPF topologies.

- PIM MT-ID: An MT-ID is a numerical identifier associated with an RPF topology.

- PIM MT-ID Join Attribute: This is a new type of Join Attribute that is introduced by this document in order to specify RPF topology in the PIM Join messages.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Functional Overview

PIM relies on routes learned from routing protocols for the purpose of RPF. These routes form one or more topologies. This section describes the function of multi-topology routing for PIM and its applicability.

3.1. PIM RPF Topology

PIM RPF topology is a collection of routes used by PIM to perform the RPF operation when building shared or source trees. The routes in the topology may be contributed by different protocols. In the rest of the document, PIM RPF topology may be simply referred to as "topology" when there is no ambiguity.

In a multi-topology environment, multiple RPF topologies can be created in the same network. A particular source may be reachable in only one of the topologies or in several of them via different paths.

To help explain the relationship between an MT-capable unicast routing protocol and MT-capable RPF topologies, consider the following example described by Figure 1.

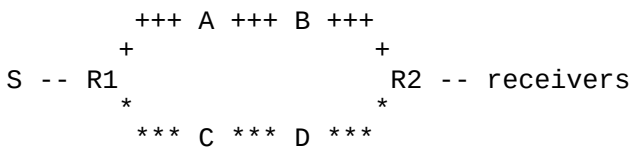


Figure 1. A simple topology for multicast

- The traffic source is S. S is announced by R1 using Multiprotocol BGP (MBGP) to every router. This route is installed in every topology.

- Two topologies are created in the unicast IGP, let us call them OSPF 1000 and OSPF 2000. OSPF 1000 includes A, B, and interfaces in R1 and R2 that are configured to be part of OSPF 1000. OSPF 2000 includes C, D, and interfaces on R1 and R2 that are configured to be part of OSPF 2000.

- Two PIM RPF topologies are created, let us call them PIM 500 and PIM 600.

PIM 500 comprises the following routes: S announced by MBGP and those learned via OSPF 1000.

PIM 600 comprises the following routes: S announced by MBGP and those learned via OSPF 2000

The above example illustrates that the naming spaces of MT-ID are not required to be the same between PIM and IGPs. Furthermore, a unicast IGP topology and the PIM RPF topology to which the IGP topology contributes routes are not required to have the same set of routes. In the above example, the prefix covering S does not exist in either OSPF 1000 or OSPF 2000, but since it exists in PIM 500 and PIM 600, R2 is able to join to it via either path.

There are two methods to select the RPF topology for a particular multicast distribution tree, via configuration or via PIM.

When it is done via configuration, a network administrator configures a policy that maps a group range to a topology and/or maps a source prefix range to a topology. Using the same example, the policy can say that to build a forwarding tree for G1 only routes in PIM 500 are to be used, and to build a forward tree for G2 only routes in PIM 600 are used. The result is that packets for (S, G1) will follow the path of S-R1-A-B-R2 and packets for (S, G2) will follow the path of S-R1-C-D-R2.

An alternative to static configuration is to include the RPF topology information as a new PIM Join Attribute in the PIM Join packets sent by downstream routers.

Both methods can be used at the same time. The details of the first method are implementation specific and are not discussed in this document. The specification to support the second method is included in this document.

3.2. PIM MT-ID

For each PIM RPF topology created, a unique numerical ID is assigned per PIM domain. This ID is called the PIM MT-ID. The PIM MT-ID has the following properties.

- It is the path identifier that is used by the PIM control plane, but it does not function in the forwarding state for a specific topology. The differentiation for topologies on the forwarding plane is made by different group addresses and/or source addresses instead.

- As shown earlier, this value is not required to be the same as the MT-ID used by the unicast routing protocols that contribute routes to the topology. In practice, when only one unicast routing protocol (such as OSPF or IS-IS) is used, the PIM MT-ID is RECOMMENDED to be assigned using the same value as the IGP topology identifier. Using the same example presented earlier, if every route in PIM 500 is contributed by OSPF 1000, it is RECOMMENDED to name this RPF topology as PIM 1000 instead of PIM 500. This is for the purpose of reducing management overhead and simplifying troubleshooting.

- This value MUST be unique and consistent within the network for the same topology. For example, PIM 500 MUST refer to the same topology on routers R1, C, D, and R2. For actual deployment, one should have a means to detect inconsistency of the PIM MT-ID configuration, but the detail of such mechanism is beyond the scope of this document.

- 0 is reserved as the default, and it MUST NOT be included in the Join Attribute encoding.

- How to assign a PIM MT-ID to a topology is decided by the network administrator and is outside the scope of this document.

3.3. Applicability

The PIM MT-ID Join Attribute described in this document applies to PIM Join/Assert packets used by PIM SM/SSM/Bidir (Sparse Mode/Source-Specific Mode/Bidirectional). It is not used in any other PIM packets. As such, it can only be used to build shared or source trees for PIM SM/SSM and PIM-Bidir downstream.

When this attribute is used in combination with RPF vectors defined in [RFC5496] and [MVPN], the vectors are processed against the topology identified by the PIM MT-ID attribute.

4. Protocol Specification of PIM MT-ID

The change to the PIM protocol includes two pieces: the PIM MT-ID Hello Option and the PIM MT-ID Join Attribute.

4.1. PIM MT-ID Hello Option

The PIM MT-ID Hello Option is used by a router to indicate if it supports the functionality described by this document. If it does, it MUST include the PIM Hello Option in its PIM Hello packets and MUST include both the Join Attribute Option [RFC5384] and the new PIM MT-ID Option (see Section 5.1 of this document for packet format).

4.2. PIM MT-ID Join Attribute

4.2.1. Sending PIM MT-ID Join Attribute

When a PIM router originates a PIM Join/Assert packet, it may choose to encode the PIM MT-ID of the topology in which RPF lookup is to take place for the corresponding (*,G) or (S,G) entry. The PIM MT-ID identifies the topology chosen by local policy/configuration or is the value received from downstream routers after MT-ID conflict resolution procedures have been applied (See Section 4.2.4 for further detail).

The following are the exceptions:

- A router SHOULD NOT include the attribute if PIM MT-ID is 0. The value of 0 is ignored on reception.

- A router SHOULD NOT include the PIM MT-ID in its Join/Assert packets if the upstream router, or any of the routers on the LAN, does not include the "PIM Join Attribute" or "PIM MT-ID" option in its Hello packets.

- A router SHOULD NOT attach PIM MT-ID for pruned sources. PIM MT-ID MUST be ignored for a pruned source by a router processing the Prune message.

4.2.2. Receiving PIM MT-ID Join Attribute

When a PIM router receives a PIM MT-ID Join Attribute in a Join/Assert packet, it MUST perform the following:

- Validate the attribute encoding. The detail is described in the next section.

- If the Join Attribute is valid, use the rules described in the section "Conflict Resolution" to determine a PIM MT-ID to use.

- Use the topology identified by the selected PIM MT-ID to perform RPF lookup for the (*,G)/(S,G) entry unless a different topology is specified by a local configuration. The local configuration always takes precedence.

While it is an exception case, it is worthwhile to describe what will happen if a router receives PIM MT-ID Join Attribute but doesn't support the functionality described in [RFC5384] or this document. If the router supports [RFC5384] but not this document, it is able to skip the PIM MT-ID Join Attribute and move on to the next Join Attribute, if one is present. The RPF decision will not be altered because the router doesn't understand the meaning of the PIM MT-ID Join Attribute. The router will use the procedures described by [RFC5384] to perform conflict resolution.

If a router doesn't support [RFC5384], it will ignore the Join/Assert message because it is not able to parse the encoded sources.

If a router does support both [RFC5384] and this document, but chooses not to send either the PIM MT-ID or the PIM Join Attribute Option in its Hello packets (likely due to administrative reasons), it SHOULD ignore the Join/Assert message when it receives a PIM Join/Assert packet with the PIM MT-ID Join Attribute.

4.2.3. Validating PIM MT-ID Join Attribute

An upstream router MUST be known to support this document in order for a downstream router to include the PIM MT-ID attribute in its Join packets. However, an upstream router doesn't need to know whether or not a downstream router supports this document when deciding whether to accept the attribute. Hence, if the Join packet sender doesn't include the "PIM Join Attribute" or "PIM MT-ID"

options in its Hello packets, the PIM MT-ID attribute in the Join may still be considered valid. This is also in accordance with the "Robustness Principle" outlined in [RFC793].

The following text specifies the detail of the validity check.

- There is at most 1 PIM MT-ID attribute encoded. If there are multiple PIM MT-ID Join Attributes included (possibly due to an error in the implementation), only the last one is accepted for this particular source. Processing of the rest of the Join message continues.

- The Length field must be 2. If the Length field is not 2, the rest of the Join message, including the current (S,G) or (*,G) entry, MUST be ignored. The group, source, and Rendezvous Point (RP) in the Join message that have already been processed SHOULD still be considered valid.

- The value MUST NOT be 0. If it is 0, the PIM MT-ID attribute is ignored. Processing of the rest of the Join message, including the current (S,G) or (*,G) entry, continues as if the particular PIM MT-ID attribute weren’t present in the packet.

4.2.4. Conflict Resolution

The definition of "PIM MT-ID conflict" varies depending on whether it is on an upstream or a downstream router.

PIM MT-ID conflicts arises on an upstream router when the router doesn’t have a local topology selection policy and receives Join packets from downstream routers and/or Assert packets from other forwarding routers on the LAN and those packets contain different PIM MT-IDs.

However, if an upstream router has a local configuration that specifies PIM MT-IDs to identify RPF topologies, and those MT-IDs do not match the MT-ID on a received Join or Assert packet, this is not considered to be a conflict and the resolution procedures are not applied. This includes the case where there are local PIM MT-IDs, but there is no PIM MT-ID encoded in the incoming packet.

On the other hand, when a downstream router sees a different PIM MT-ID attribute from other routers on the LAN, it applies rules to resolve the conflicts regardless of whether or not the router has local topology selection policy.

When two PIM MT-IDs are compared, only the 12-bit Value field (see Section 5.2) is compared. Other fields of the PIM MT-ID Join Attribute TLV Format (including the four reserved bits) MUST NOT be used in the comparison.

4.2.4.1. Conflict Resolution Rules for Upstream Routers

- If an upstream router receives different PIM MT-ID attributes from PIM Join packets, it MUST follow the rules specified in [RFC5384] to select one. The PIM MT-ID chosen will be the one encoded for its upstream neighbor.

  In order to minimize the chances of potential transient forwarding loops, an upstream router MAY choose to ignore the incoming PIM Join packets altogether if it sees a conflict in PIM MT-ID attributes. This action may also be taken by an upstream router that has locally configured topology selection policy, as an exception to the rules described above.

- If an upstream router receives a different PIM MT-ID attribute in an Assert packet, it MUST use the tiebreaker rules as specified in [RFC4601] to determine an Assert winner. PIM MT-ID is not considered in deciding a winner from Assert process.

4.2.4.2. Conflict Resolution Rules for Downstream Routers

- If a downstream router sees different PIM MT-ID attributes from PIM Join packets, it MUST follow the specification of [RFC4601] as if the attribute did not exist. For example, the router suppresses its own Join packet if a Join for the same (S,G) is seen.

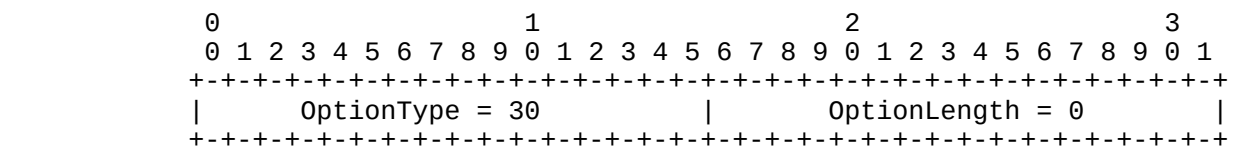
  The router MUST NOT use the rules specified in [RFC5384] to select a PIM MT-ID from Join packets sent by other downstream routers.

- If a downstream router sees its preferred upstream router loses in the Assert process, and the Assert winner uses a different PIM MT-ID, the downstream router SHOULD still choose the Assert winner as the RPF neighbour, but it MUST NOT encode PIM MT-ID when sending Join packets to it.

5. Packet Format

This section describes the format of new PIM messages introduced by this document. The messages follow the same transmission order as the messages defined in [RFC4601].

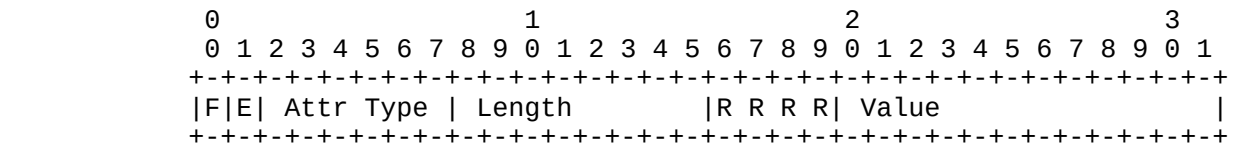
5.1. PIM MT-ID Hello Option



- OptionType: 30.

- OptionLength: 0.

5.2. PIM MT-ID Join Attribute TLV Format



- F bit: 0 Non-transitive Attribute.

- E bit: As specified by [RFC5384].

- Attr Type: 2

- Length: 2.

- R: Reserved bits, 4 in total. Set to zero on transmission. Ignored upon receipt.

- Value: PIM MT-ID, 1 to 4095.

6. IANA Considerations

6.1. PIM MT-ID Hello Option

IANA maintains a registry of "Protocol Independent Multicast (PIM) Parameters" with a sub-registry called "PIM-Hello Options".

The IANA has assigned the PIM Hello Option type value 30 for the PIM MT-ID Hello Option according to the First Come First Served allocation policy.

The IANA has assigned a Length value of 0.

6.2. PIM MT-ID Join Attribute Type

The IANA maintains a registry of "Protocol Independent Multicast (PIM) Parameters" with a sub-registry called "PIM Join Attribute Types".

The IANA has assigned a value of 2 for the PIM MT-ID Join Attribute defined in Section 5.2 of this document.

7. Security Considerations

As described in [RFC5384], the security of the Join Attribute is only guaranteed by the security of the PIM packet that carries it. Similarly, the security of the Hello Option is only guaranteed by securing the whole Hello Packet.

In view of the fact that malicious alteration of the PIM MT-ID Hello Option or the PIM MT-ID carried in a packet might cause the PIM resiliency goals to be violated, the security considerations of [RFC4601] apply to the extensions described in this document.

As a type of PIM Join Attribute, the security considerations described in [RFC5384] apply here. Specifically, malicious alteration of PIM MT-ID may cause the resiliency goals to be violated.

8. Acknowledgments

The authors would like to thank Eric Rosen, Ice Wijnands, Dino Farinacci, Colby Barth, Les Ginsberg, Dimitri Papadimitriou, Thomas Morin, and Hui Liu for their input.

The authors would also like to thank Adrian Farrel for his detailed and constructive comments during the AD review.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

[RFC5384] Boers, A., Wijnands, I., and E. Rosen, "The Protocol Independent Multicast (PIM) Join Attribute Format", RFC 5384, November 2008.

9.2. Informative References

[RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, June 2007.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC5496] Wijnands, IJ., Boers, A., and E. Rosen, "The Reverse Path Forwarding (RPF) Vector TLV", RFC 5496, March 2009.

[MVPN] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", Work in Progress, January 2010.

Authors' Addresses

Yiqun Cai
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134

EMail: ycai@cisco.com

Heidi Ou
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134

EMail: hou@cisco.com