Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 9, 2015

J. Asghar
IJ. Wijnands, Ed.
S. Krishnaswamy
A. Karan
Cisco Systems
V. Arya
DIRECTV Inc.

February 5, 2015

Explicit RPF Vector
draft-ietf-pim-explicit-rpf-vector-06.txt

Abstract

The PIM Reverse Path Forwarding (RPF) Vector TLV defined in RFC 5496 can be included in a PIM Join Attribute such that the RPF neighbor is selected based on the unicast reachability of the RPF Vector instead of the Source or RP associated with the multicast tree.

This document defines a new RPF Vector Attribute type such that an explicit RPF neighbor list can be encoded in the PIM Join Attribute, bypassing the unicast route lookup.

Table of Contents

1. Introduction

The procedures in [RFC5496] define how a RPF vector can be used to influence the path selection in the absence of a route to the source. The same procedures can be used to override a route to the source when it exists. It is possible to include multiple RPF vectors in the list where each router along the path will perform a unicast route lookup on the first vector in the attribute list. Once the router owning the address of the RPF vector is reached, following the procedures in [RFC5496], the RPF vector will be removed from the attribute list. This will result in a 'loosely' routed path based on the unicast reachability of the RPF vector(s). We call this 'loosely' because we still depend on unicast routing reachability to the RPF Vector.

In some scenarios we don't want to rely on the unicast reachability to the RPF vector address and we want to build a path strictly based on the RPF vectors. In that case the RPF vectors represent a list of directly connected PIM neighbors along the path. For these vectors

we MUST NOT do a unicast route lookup. We call these 'Explicit' RPF Vector addresses. If a router receiving an Explicit RPF Vector does not have a PIM neighbor matching the Explicit RPF Vector address it MUST NOT fall back to loosely routing the join. Instead, it may process the packet and store the RPF Vector list so that the PIM join may be sent out as soon as the neighbor comes up. Since the behavior of the Explicit RPF Vector differs from the loose RPF vector as defined [RFC5496], we're defining a new attribute called the Explicit RPF Vector.

This document defines a new TLV in the PIM Join Attribute message [RFC5384] for specifying the explicit path.

2. Specification of Requirements

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

3. Motivation

Some broadcast video transport networks use a multicast PIM Live-Live resiliency model for video delivery based on PIM SSM or PIM ASM. Live-Live implies using 2 active spatially diverse multicast trees to

transport video flows from root to leaf multicast routers. The leaf multicast router receives 2 copies from the PIM multicast core and will replicate 1 copy towards the receivers [I-D.ietf-rtgwg-mofrr].

One of the requirements of the PIM Live-Live resiliency model is to ensure path-diversity of the 2 active PIM trees in the core such that they do not intersect to avoid a single point of failure. IGP routed RPF paths of 2 PIM trees could be routed over the same transit router and create a single point of failure. It is useful to have a way to specify the explicit path along which the PIM join is propagated.

How the Explicit RPF Vector list is determined is outside the scope of this document. For example, it may either be manually configured by the network operator or procedures may be implemented on the egress router to dynamically calculate the vector list based on a link state database protocol, like OSPF or IS-IS.

Due to the fact that the leaf router receives two copies of the multicast stream via two diverse paths, there is no need for PIM to repair the broken path immediately. It is up to the egress router to either wait for the broken path to be repaired or build a new explicit path using a new RPF vector list. Which method is applied depends very much on how the vector list was determined initially.

Double failures are not considered and are outside the scope of this document.

This document describes the procedures to carry Explicit RPF vectors in PIM, but it does not introduce any new mechanism in PIM to validate the correctness of the RPF vectors. It is up to the mechanism(s) that produce the Explicit RPF Vectors to ensure they are correct. Existing mechanisms like [I-D.ietf-mboned-mtrace-v2] may be used to verify how the PIM tree was build.

4. Use of the PIM Explicit RPF Vector

Figure 1 provides an example multicast join path R4->R3->R6->R5->R2->R1, where the multicast join is explicitly routed to the source hop-by-hop using the Explicit RPF Vector list. When R5-R6 link fails the join will NOT take an alternate path.

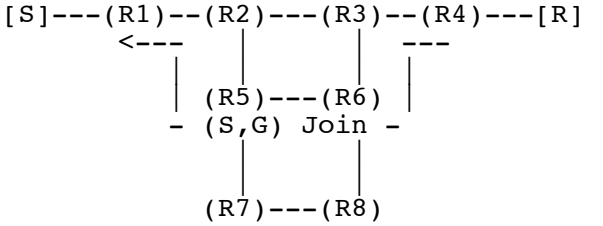

Figure 1

In comparison, when [RFC5496] procedures are used, if R5-R6 link fails then the join may be re-routed using R6-R8-R7 path to reach R5.

5. Explicit RPF Vector Attribute

This draft uses PIM join attribute type TBD1 by IANA for specifying an Explicit RPF Vector.

6. Mixed Vector Processing

Explicit RPF Vector attribute does not impact or restrict the functionality of other RPF vector attributes in a PIM join. It is possible to mix vectors of different types, such that some part of the tree is explicit and other parts are loosely routed. RPF vectors are processed in the order in which they are specified.

7. Conflicting RPF Vectors

It is possible that a PIM router has multiple downstream neighbors. If for the same multicast route there is an inconsistency between the Explicit RPF Vector lists provided by the downstream PIM neighbor, the procedures as documented in section 3.3.3 [RFC5384] apply.

8. PIM Asserts

Section 3.3.3 of [RFC5496] specifies the procedures for how to deal with PIM asserts when RPF vectors are used. The same procedures apply to the Explicit RPF Vector. There is minor behavioral difference, the route metric that is included in the PIM Assert should be the route metric of the first Explicit RPF vector address in the list. However, the first Explicit vector should always be directly connected, so the Metric may likely be zero. The Metric will therefore not be a tie breaker in the PIM Assert selection procedure.

9. Join Suppression

Section 3.3.4 of [RFC5496] specifies the procedures how to apply join suppression when an RPF Vector attribute is included in the PIM join. The same procedure applies to the Explicit RPF Vector attribute. The procedure MUST match against all the Explicit RPF Vectors in the PIM join before a PIM join can be suppressed.

10. Unsupported Explicit Vector Handling

The F bit MUST be set to 0 in all Explicit RPF vectors in case the upstream router receiving the join does not support the TLV. As described in section 3.3.2 of [RFC5384], routers that do not understand the type of a particular attribute that has the F bit clear will discard it and continue to process the join.

This processing is particularly important when the routers that do not support the Explicit RPF TLV are identified as hops in the explicit RPF list, because failing to remove the RPF vectors could cause upstream routers to send the join back toward these routers causing loops.

11. Explicit RPF Vector Attribute TLV Format

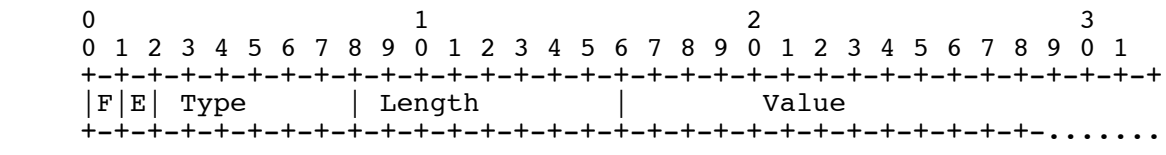

Figure 2

F bit: The F bit MUST be set to 0. Otherwise there could be loops.

E bit: End of Attributes. If this bit is set then this is the last TLV specified in the list.

Type: The Vector Attribute type is TBD1.

Length: Length depending on the Address Family of the Encoded-Unicast address.

Value: Encoded-Unicast address. This could be a valid primary or secondary address.

12. IANA Considerations

A new attribute (TBD1) type from the "PIM Join Attribute Types" registry needs to be assigned by IANA for the Explicit RPF Vector attribute. The proposed value 4.

13. Security Considerations

Security of the Explicit RPF Vector Attribute is only guaranteed by the security of the PIM packet, so the security considerations for PIM Join packets as described in PIM-SM [RFC4601] apply here. Additionally, the Explicit RPF Vector list should be subject to a policy to validate the list consists of a valid path before its used by a receiver to build a multicast tree.

14. Acknowledgments

The authors would like to thank Vatsa Kumar, Nagendra Kumar and Bharat Joshi for the comments on the document.

15. References

15.1. Normative References

[I-D.ietf-rtgwg-mofrr]
Karan, A., Filsfils, C., Wijnands, I., and B. Decraene, "Multicast only Fast Re-Route", draft-ietf-rtgwg-mofrr-05 (work in progress), February 2015.

[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.

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

15.2. Informative References

[I-D.ietf-mboned-mtrace-v2]
Asaeda, H. and W. Lee, "Mtrace Version 2: Traceroute Facility for IP Multicast", draft-ietf-mboned-mtrace-v2-11 (work in progress), October 2014.

Authors' Addresses

Javed Asghar
Cisco Systems
725, Alder Drive
Milpitas CA 95035
USA

Email: jasghar@cisco.com

IJsbrand Wijnands (editor)
Cisco Systems
De Kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Sowmya Krishnaswamy
Cisco Systems
3750 Cisco Way
San Jose CA 95134
USA

Email: sowkrish@cisco.com

Apoorva Karan
Cisco Systems
3750 Cisco Way
San Jose CA 95134
USA

Email: apoorva@cisco.com

Vishal Arya
DIRECTV Inc.
2230 E Imperial Hwy
El Segundo CA 90245
USA

Email: varya@directv.com