Network Working Group
Internet-Draft
Intended status: Experimental
Expires: December 16, 2016

J. Arango
S. Venaas
Cisco Systems
I. Kouvelas
Arista Networks Inc.
D. Farinacci
lispers.net

June 14, 2016

PIM Join Attributes for LISP Environments
draft-ietf-pim-join-attributes-for-lisp-04.txt

Abstract

This document defines two PIM Join/Prune attributes that support the construction of multicast distribution trees where the root and receivers are located in different LISP sites. These attributes allow the receiver site to select between unicast and multicast underlay transport and to convey the receiver ETR's RLOC address to the control plane of the root ITR.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on December 16, 2016.

Copyright Notice

Table of Contents

1. Introduction

The construction of multicast distribution trees where the root and receivers are located in different LISP sites [RFC6830] is defined in [RFC6831]. Creation of (root-EID,G) state in the root site requires that unicast LISP-encapsulated Join/Prune messages be sent from an ETR on the receiver site to an ITR on the root site.

[RFC6831] specifies that (root-EID,G) data packets are to be LISP-encapsulated into (root-RLOC,G) multicast packets. However, a wide deployment of multicast connectivity between LISP sites is unlikely to happen any time soon. In fact, some implementations are initially focusing on unicast transport with head-end replication between root and receiver sites.

The unicast LISP-encapsulated Join/Prune message specifies the (root-EID,G) state that needs to be established in the root site, but conveys nothing about the receivers capability or desire to use multicast as the underlying transport. This document specifies a Join/Prune attribute that allows the receiver ETR to select the desired transport.

Knowledge of the receiver ETR's RLOC address is also essential to the control plane of the root ITR. It determines the downstream destination for unicast head-end replication and identifies the

receiver ETR that needs to be notified should the root of the distribution tree move to another site.

Service providers may implement URPF policies requiring that the outer source address of the LISP-encapsulated Join/Prune message be the address of the receiver ETR's core-facing interface used to physically transmit the message. However, due to policy and load balancing considerations, the outer source address may not be the RLOC on which the receiver site wishes to receive a particular flow. This document specifies a Join/Prune attribute that conveys the appropriate receiver ETR's RLOC address to the control plane of the root ITR.

2. Requirements Notation

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. PIM Join/Prune Attributes

PIM Join/Prune attributes are defined in [RFC5384] by introducing a new Encoded-Source type that, in addition to the Join/Prune source, can carry multiple type-length-value (TLV) attributes. These attributes apply to the individual Join/Prune sources on which they are stored.

The attributes defined in this document conform to the format of the encoding type defined in [RFC5384]. The attributes would typically be the same for all the sources in the Join/Prune message. Hence we RECOMMEND using the hierarchical Join/Prune attribute scheme defined in [I-D.ietf-pim-hierarchicaljoinattr]. This hirarchichal system allows attributes to be conveyed on the Upstream Neighbor Address field, thus enabling the efficient application of a single attribute instance to all the sources in the Join/Prune message.

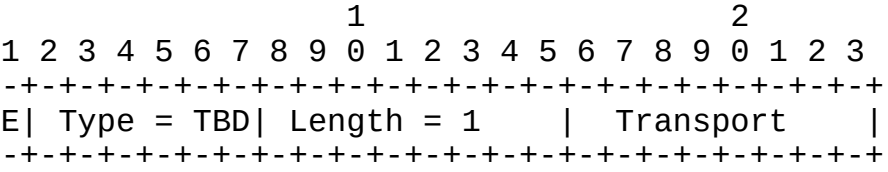
LISP xTRs do not exchange PIM Hello Messages and hence no Hello option is defined to negotiate support for these attributes. Systems that support unicast head-end replication are assumed to support these attributes.

4. The Transport Attribute

It is essential that a mechanism be provided by which the desired transport can be conveyed by receiver sites. Root sites with multicast connectivity will want to leverage multicast replication. However, not all receiver sites can be expected to have multicast connectivity. It is thus desirable that root sites be prepared to

support (root-EID,G) state with a mixture of multicast and unicast output state. This document specifies a Join/Prune attribute that allows the receiver to select the desired underlying transport.

4.1. Transport Attribute Format



F-bit: The Transitive bit. Specifies whether the attribute is transitive or non-transitive. MUST be set to zero. This attribute is ALWAYS non-transitive.

E-bit: End-of-Attributes bit. Specifies whether this attribute is the last. Set to zero if there are more attributes. Set to 1 if this is the last attribute.

Type: The Transport Attribute type is TBD.

Length: The length of the Transport Attribute value. MUST be set to 1.

Transport: The type of transport being requested. Set to 0 for multicast. Set to 1 for unicast.

4.2. Using the Transport Attribute

Hierarchical Join/Prune attribute instances [I-D.ietf-pim-hierarchicaljoinattr] SHOULD be used when the same Transport Attribute is to be applied to all the sources within the Join/Prune message or all the sources within a group set. The root ITR MUST accept Transport Attributes in the Upstream Neighbor Encoded-Unicast address, Encoded-Group addresses, and Encoded-Source addresses.

There MUST NOT be more than one Transport Attribute within the same encoded address. If an encoded address has more than one instance of the attribute, the root ITR MUST discard all affected Join/Prune sources.

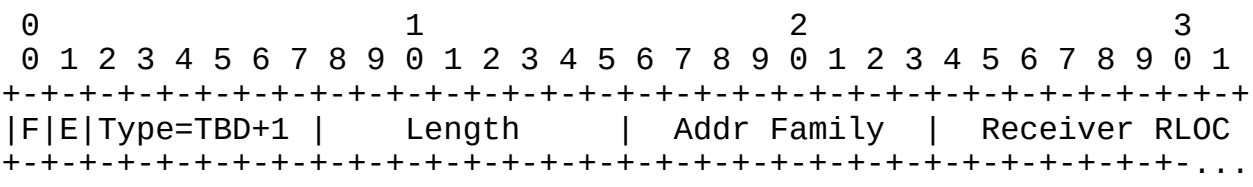
5. Receiver ETR RLOC Attribute

When a receiver ETR requests unicast head-end replication for a given (root-EID,G) entry, the PIM control plane of the root ITR must maintain an output interface list ("oif-list") entry for the receiver ETR and its corresponding RLOC address. This allows the root ITR to perform unicast LISP-encapsulation of multicast data packets to each and every receiver ETR that has requested unicast head-end replication.

The PIM control plane of the root ITR could potentially determine the RLOC address of the receiver ETR from the outer source address field of LISP-encapsulated Join/Prune message. However, receiver ETRs are subject to URPF checks by the network providers on each core-facing interface. The outer source address must therefore be the RLOC of the core-facing interface used to physically transmit the LISP-encapsulated Join/Prune message. Due to policy and load balancing considerations, that may not be the RLOC on which the receiver site wishes to receive a particular flow. This document specifies a Join/Prune attribute that conveys the appropriate receiver RLOC address to the PIM control plane of the root ITR.

To support root-EID mobility, receiver ETRs must also be tracked by the LISP control plane of the root ITR, regardless of the underlying transport. When the root-EID moves to a new root ITR in a different LISP site, the receiver ETRs do not know the root-EID has moved and therefore do not know the RLOC of the new root ITR. This is true for both unicast and multicast transport modes. The new root ITR does not have any receiver ETR state. Therefore, it is the responsability of the old root ITR to inform the receiver ETRs that the root-EID has moved. When the old root ITR detects that the root-EID has moved, it sends a LISP SMR message to each receiver ETR. The receiver ETRs do a mapping database lookup to retrieve the RLOC of the new root ITR. The old root ITR detects that the root-EID has moved when it receives a Map-Notify from the Map-Server. The transmission of the Map-Notify is triggered when the new root ITR registers the root-EID [I-D.portoles-lisp-eid-mobility]. When a receiver ETR determines that the root ITR has changed it will send a LISP-encapsulated PIM prune message to the old root XTR and a LISP-encapsulated PIM join message to the new root XTR.

5.1. Receiver RLOC Attribute Format



F-bit: The Transitive bit. Specifies whether this attribute is transitive or non-transitive. MUST be set to zero. This attribute is ALWAYS non-transitive.

E-bit: End-of-Attributes bit. Specifies whether this attribute is the last. Set to zero if there are more attributes. Set to 1 if this is the last attribute.

Type: The Receiver RLOC Attribute type is TBD+1.

Length: The length in octets of the attribute value. MUST be set to the length in octets of the receiver RLOC address plus one octet to account for the Address Family field.

Addr Family: The PIM Address Family of the receiver RLOC as defined in [RFC4601].

Receiver RLOC: The RLOC address on which the receiver ETR wishes to receiver the unicast-encapsulated flow.">

5.2. Using the Receiver RLOC Attribute

Hierarchical Join/Prune attribute instances [I-D.ietf-pim-hierarchicaljoinattr] SHOULD be used when the same Receiver RLOC attribute is to be applied to all the sources within the message or all the sources within a group set. The root ITR MUST accept Transport Attributes in the Upstream Neighbor Encoded-Unicast address, Encoded-Group addresses, and Encoded-Source addresses.

There MUST NOT be more than one Receiver RLOC Attribute within the same encoded address. If an encoded address has more than one instance of the attribute, the root ITR MUST discard all affected Join/Prune sources.

6. Security Considerations

Security of the Join Attribute is only guaranteed by the security of the PIM packet. The attributes specified herein do not enhance or diminish the privacy or authenticity of a Join/Prune message. A site that legitimately or maliciously sends and delivers a Join/Prune

message to another site will equally be able to append these and any other attributes it wishes.

7. IANA Considerations

Two new PIM Join/Prune attribute types need to be assigned. Type 5 is being requested for the Transport Attribute. Type 6 is being requested for the Receiver RLOC Attribute.

8. Normative References

[AFI] IANA, , "Address Family Numbers", http://www.iana.org/assignments/address-family-numbers.

[I-D.ietf-pim-hierarchicaljoinattr] Venaas, S., Arango, J., and I. Kouvelas, "Hierarchical Join/Prune Attributes", draft-ietf-pim-hierarchicaljoinattr-08 (work in progress), April 2016.

[I-D.portoles-lisp-eid-mobility] Portoles-Comeras, M., Ashtaputre, V., Moreno, V., Maino, F., and D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified Control Plane", draft-portoles-lisp-eid-mobility-00 (work in progress), April 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, DOI 10.17487/RFC4601, August 2006, <http://www.rfc-editor.org/info/rfc4601>.

[RFC5384] Boers, A., Wijnands, I., and E. Rosen, "The Protocol Independent Multicast (PIM) Join Attribute Format", RFC 5384, DOI 10.17487/RFC5384, November 2008, <http://www.rfc-editor.org/info/rfc5384>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <http://www.rfc-editor.org/info/rfc6830>.

[RFC6831] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas, "The Locator/ID Separation Protocol (LISP) for Multicast Environments", RFC 6831, DOI 10.17487/RFC6831, January 2013, <http://www.rfc-editor.org/info/rfc6831>.

Authors' Addresses

Jesus Arango
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: jearango@cisco.com

Stig Venaas
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: stig@cisco.com

Isidor Kouvelas
Arista Networks Inc.
5453 Great America Parkway
Santa Clara, CA 95054
USA

Email: kouvelas@arista.com

Dino Farinacci
lispers.net
San Jose, CA
USA

Email: farinacci@gmail.com