

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: September 4, 2017

J. Arango
J. Leong
Cisco Systems
March 3, 2017

LISP Stateful Pull Model
draft-jearango-lisp-stateful-pull-model-00.txt

Abstract

This document specifies a stateful pull model for LISP where ITRs can subscribe with the mapping system to be notified whenever a particular EID mapping changes. The model uses a publish/subscribe mechanism that supports overlapping EID registrations without having to notify the ITR about every single prefix covered by a particular subscription. The pull model is stateful in the sense that it requires that the mapping system maintain subscription state.

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 September 4, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 2
 2. Requirements Notation 2
 3. Subscription Establishment 2
 4. Publication of Overlapping Prefixes 3
 5. Mobility and Barrier Prefixes 4
 6. Negative Subscriptions 4
 7. ITR Eviction of Subscription and Map-cache State 5
 8. Normative References 6
 Authors' Addresses 6

1. Introduction

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. Subscription Establishment

When operating in a stateful pull model, a lookup miss in the data plane's forwarding table results in the transmission of a subscription message to the mapping system. The subscription message contains a query prefix field that is set to the destination host's address. As in the stateless pull model, the ITR creates an incomplete map-cache entry to inhibit further signaling from the data plane until the corresponding mapping information is received from the mapping system.

When processing a subscription message, the mapping system performs a longest-prefix match in the mapping database for the query prefix included in the subscription message. The resulting mapping is sent to the ITR in a publication message. The query prefix is also included as an explicit indication that the publication message is to be used by the ITR to create subscription state.

The ITR creates an entry in a subscription table when it receives a publication message that includes a query prefix. The subscription entry contains the longest-matching EID prefix returned by the mapping system in the publication message. The query prefix is used by the ITR to perform an exact match lookup for an incomplete map-cache entry. The map-cache entry is then linked as one of the sources (producers) of the subscription state. The host address of

the map-cache entry is overwritten with the EID prefix of the subscription entry. The locators of the subscription entry are copied to the map-cache entry. The map-cache entry is transitioned to complete state and reprogramed in the data plane.

4. Publication of Overlapping Prefixes

The mapping system must publish all mappings that are more specific than the subscription state, unless a mechanism is conceived to more efficiently support overlapping prefixes. Publication of more specific mappings is necessary to ensure that the correct locators are used in the following two situations:

- a. An ITR can quite possibly have a subscription entry that covers multiple destination hosts to which it is sending traffic. The mapping system could later receive a registration that is more specific than the current subscription state and at the same time happens to be the longest matching prefix for one or multiple destination hosts currently being covered by the subscription.
- b. An ITR starts forwarding traffic to a new destination host that is covered by an existing subscription entry, but for which the mapping system has a mapping that is even more specific.

To efficiently support overlapping prefixes, the mapping system only publishes registered mappings that are one level below the subscription prefix. To be formally precise, the mapping system only publishes registered mappings that are direct hierarchical children of the subscription prefix. A more specific registered mapping is a hierarchical child of a subscription prefix if there DOES NOT exist another registered mapping that is also more specific than the subscription prefix and at the same time less specific than the prefix under consideration.

The publication messages for child prefixes have a flag set indicating that a corresponding map-cache entry should be created with a "signal-and-forward" action in the data plane. These Map-cache entries are also marked or linked as additional sources (producers) of the corresponding subscription. When a longest matching forwarding lookup hits one of these map-cache entries, the packet will be encapsulated using one of the locators in the entry and the control plane will be signaled to trigger a subscription for the destination host, as described in section Section 3. The ITR is now subscribed to what used to be a child prefix. The "signal-and-forward" map-cache entry is deleted. The incomplete map-cache entry created during the data-plane signaling event inherits the EID prefix and the locators from the new subscription state. Any registered

mappings that happen to be hierarchical children of the new subscription will be published by the mapping system.

5. Mobility and Barrier Prefixes

Dynamic EID prefixes are mobility prefixes configured at the ETRs to enable detection and registration of individual hosts covered by the prefix, as opposed to just registering the entire block as a single prefix. Depending on the prefix length, dynamic EID prefixes can result in a very large fan-out of individual host mappings that are hierarchical children of a less specific mapping.

An ITR that sends a subscription where the query prefix is a mobile host that is not yet registered in the mapping system can potentially end up being subscribed to a less specific prefix with a huge fan-out of mobile host mappings. These mobile host mappings are hierarchical children of the less specific mapping and will therefore be published to the ITR.

To avoid publishing all the mobile hosts to the ITR, the dynamic EID prefix is configured in the mapping system as a barrier prefix. A barrier prefix has the following properties:

- a. If the barrier prefix is the longest matching prefix for a subscription request, the ITR gets subscribed to the query prefix, not the barrier prefix. If the mobile node corresponding to the query prefix is not yet registered, the publication message can have an empty locator set.
- b. A barrier prefix that is a hierarchical child of a less specific subscription gets published to the ITR with an explicit indication that it **MUST** be programmed in the data plane with a "signal" action. This ensures that future traffic covered by the dynamic EID prefix triggers a subscription, even if the ITR has a map-cache entry that is less specific than the dynamic EID prefix.

6. Negative Subscriptions

In the LISP stateless pull model, a Map-Request whose longest matching lookup in the mapping system results in a lookup miss will trigger a negative Map-Reply. The prefix included in the reply is the least specific gap or hole in the mapping system that also covers the prefix in the Map-Request. This minimizes the amount of map-cache entries necessary to forward traffic not covered by mapping database.

In the stateful pull model, a subscription message could similarly result in a lookup miss in the mapping system. The least specific gap or hole in the mapping database that covers the subscription request gets self-registered by the mapping system as a negative mapping. A negative mapping can contain an empty locator set. Alternatively, if one or more PETRs are registering the 0/0 prefix, the locator set of a negative mappings could inherit the merged set of locators from the 0/0 mapping. The publication message sent to the ITR in response to the subscription contains the negative mapping. According to the subscription procedure described in section Section 3, the ITR creates a subscription entry and a map-cache entry for the negative mapping.

If and when a prefix gets registered that is more specific than a negative mapping, said prefix will effectively be a hierarchical child of the negative mapping and will therefore be published to any ITR currently subscribed to the negative prefix.

Negative mappings get evicted from the mapping system when the set of ITRs subscribed to the negative mapping becomes empty. Subscriptions to prefixes that are more specific than the negative mapping do not prevent the eviction of the negative mapping. What is important is that there are no subscriptions on the negative mapping itself.

7. ITR Eviction of Subscription and Map-cache State

An ITR subscription entry may be sourced by one map-cache entry whose prefix is equal to that of the subscription entry, and multiple map-cache entries whose prefixes are more specific than that of the subscription entry. Any of these map-cache entries may be evicted when an unpublish message is received from the mapping system. An unpublish message is an indication that the corresponding mapping is no longer registered in the mapping system.

A subscription entry can be evicted from the ITR when the set of map-cache entries sourcing the subscription becomes empty. The ITR sends an unsubscribe message when a suscription entry is evicted.

The ITR can collectively evict a subscription entry and all associated map-cache entries in a single atomic operation if none of the map-cache entries have been hit by forwarding traffic after an unspecified amount of time. Note that while evicting a single map-cache entry for lack of use is possible, it will not prevent the mapping system for publishing it again if there is a change in the corresponding mapping.

Note that the mapping system MUST keep track of a subscription even if there are no registered mappings that are covered by the

subscription. The Mapping system can only evict a subscription entry if it receives an unsubscribe message from the ITR or if it loses communication with the ITR

8. Normative References

- [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>>.
- [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>>.

Authors' Addresses

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

Email: gearango@cisco.com

Johnson Leong
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: joleong@cisco.com