Abstract

This document describes extentions to the Locator/ID Separation Protocol (LISP) Data-Plane, via changes to the LISP header, to support multi-protocol encapsulation.

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 https://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 May 22, 2020.

Copyright Notice

Copyright (c) 2019 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 (https://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
   1.1. Conventions ........................................ 3
   1.2. Definition of Terms ................................ 3
2. LISP Header Without Protocol Extensions .................. 3
3. Generic Protocol Extension for LISP (LISP-GPE) ............ 4
4. Implementation and Deployment Considerations .............. 6
   4.1. Applicability Statement .......................... 6
   4.2. Congestion Control Functionality .................. 7
   4.3. UDP Checksum ..................................... 7
   4.3.1. UDP Zero Checksum Handling with IPv6 ............ 8
   4.4. Ethernet Encapsulated Payloads .................... 9
5. Backward Compatibility ................................... 10
   5.1. Detection of ETR Capabilities .................... 10
6. IANA Considerations ..................................... 10
   6.1. LISP-GPE Next Protocol Registry .................. 10
7. Security Considerations .................................. 10
8. Acknowledgements and Contributors ........................ 11
9. References .............................................. 11
   9.1. Normative References ............................ 11
   9.2. Informative References ............................ 12
Authors’ Addresses ......................................... 13

1. Introduction

The LISP Data-Plane is defined in [I-D.ietf-lisp-rfc6830bis]. It specifies an encapsulation format that carries IPv4 or IPv6 packets (henceforth jointly referred to as IP) in a LISP header and outer UDP/IP transport.

The LISP Data-Plane header does not specify the protocol being encapsulated and therefore is currently limited to encapsulating only IP packet payloads. Other protocols, most notably Virtual eXtensible Local Area Network (VXLAN) [RFC7348] (which defines a similar header format to LISP), are used to encapsulate Layer-2 (L2) protocols such as Ethernet.

This document defines an extension for the LISP header, as defined in [I-D.ietf-lisp-rfc6830bis], to indicate the inner protocol, enabling the encapsulation of Ethernet, IP or any other desired protocol all the while ensuring compatibility with existing LISP deployments.
A flag in the LISP header, called the P-bit, is used to signal the presence of the 8-bit Next Protocol field. The Next Protocol field, when present, uses 8 bits of the field that was allocated to the echo-noncing and map-versioning features in [I-D.ietf-lisp-rfc6830bis].

Since all of the reserved bits of the LISP Data-Plane header have been allocated, LISP-GPE can also be used to extend the LISP Data-Plane header by defining Next Protocol "shim" headers that implements new data plane functions not supported in the LISP header. For example, the use of the Group-Based Policy (GBP) header [I-D.lemon-vxlan-lisp-gpe-gbp] or of the In-situ Operations, Administration, and Maintenance (IOAM) header [I-D.brockners-ippm-ioam-vxlan-gpe] with LISP-GPE, can be considered an extension to add support in the Data-Plane for Group-Based Policy functionalities or IOAM metadata.

Nonce, Map-Versioning and Locator Status Bit fields are not part of the LISP-GPE header. Shim headers can be used to specify features such as echo-noncing, map-versioning or reachability by defining fields of the same size, or larger, of those specified in [I-D.ietf-lisp-rfc6830bis].

1.1. Conventions

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

1.2. Definition of Terms

This document uses terms already defined in [I-D.ietf-lisp-rfc6830bis].

2. LISP Header Without Protocol Extensions

As described in Section 1, the LISP header has no protocol identifier that indicates the type of payload being carried. Because of this, LISP is limited to carrying IP payloads.

The LISP header [I-D.ietf-lisp-rfc6830bis] contains a series of flags (some defined, some reserved), a Nonce/Map-version field and an instance ID/Locator-status-bit field. The flags provide flexibility to define how the various fields are encoded. Notably, Flag bit 5 is the last reserved bit in the LISP header.
3. Generic Protocol Extension for LISP (LISP-GPE)

This document defines two changes to the LISP header in order to support multi-protocol encapsulation: the introduction of the P-bit and the definition of a Next Protocol field. This is shown in Figure 2 and described below.

Bits 0-3 and 8-23: Bits 0-3 and 8-23 of the LISP-GPE header are Reserved. They MUST be set to zero on transmission and ignored on receipt.

Features that were implemented with bits 0-3 and 8-23 in [I-D.ietf-lisp-rfc6830bis], such as echo-noncing, map-versioning and reachability, can be implemented by defining the appropriate shim headers.

Instance ID When the I-Bit is set to 1 the high-order 24 bits of the Instance ID field are used as an Instance ID, as specified in [I-D.ietf-lisp-rfc6830bis]. The low-order 8 bits are set to zero, as the Locator-Status-Bits feature is not supported in LISP-GPE.

P-Bit: Flag bit 5 is defined as the Next Protocol bit.
If the P-bit is clear (0) the LISP header is bit-by-bit equivalent to the definition in [I-D.ietf-lisp-rfc6830bis] with bits N, L, E and V set to 0.

The P-bit is set to 1 to indicate the presence of the 8 bit Next Protocol field. The combinations of bits that are allowed when the P-bit is set are the same allowed by [I-D.ietf-lisp-rfc6830bis] when bits N, L, E and V are set to 0.

Next Protocol: The lower 8 bits of the first 32-bit word are used to carry a Next Protocol. This Next Protocol field contains the protocol of the encapsulated payload packet.

This document defines the following Next Protocol values:

- 0x01 : IPv4
- 0x02 : IPv6
- 0x03 : Ethernet
- 0x04 : Network Service Header (NSH) [RFC8300]
- 0x05 to 0x7F: Unassigned
- 0x80 to 0xFF: Unassigned (shim headers)

The values are tracked in the IANA LISP-GPE Next Protocol Registry as described in Section 6.1.

Next protocol values from 0x80 to 0xFF are assigned to protocols encoded as generic "shim" headers. All shim protocols MUST use the header structure in Figure 3, which includes a Next Protocol field. When a shim header is used with other protocols identified by next protocol values from 0x0 to 0x7F, the shim header MUST come before the further protocol, and the next header of the shim will indicate which protocol follows the shim header.

Shim headers can be used to incrementally deploy new GPE features, keeping the processing of shim headers known to a given xTR implementation in the ‘fast’ path (typically an ASIC), while punting the processing of the remaining new GPE features to the ‘slow’ path.

Shim protocols MUST have the first 32 bits defined as:
Where:

Type: This field identifies the different messages of this protocol.

Length: The length, in 4-octect units, of this protocol message not including the first 4 octects.

Reserved: The use of this field is reserved to the protocol defined in this message.

Next Protocol Field: The next protocol field contains the protocol of the encapsulated payload. The values are tracked in the IANA LISP-GPE Next Protocol Registry as described in Section 6.1.

4. Implementation and Deployment Considerations

4.1. Applicability Statement

LISP-GPE conforms, as an UDP-based encapsulation protocol, to the UDP usage guidelines as specified in [RFC8085]. The applicability of these guidelines are dependent on the underlay IP network and the nature of the encapsulated payload.

[RFC8085] outlines two applicability scenarios for UDP applications, 1) general Internet and 2) controlled environment. The controlled environment means a single administrative domain or adjacent set of cooperating domains. A network in a controlled environment can be managed to operate under certain conditions whereas in general Internet this cannot be done. Hence requirements for a tunnel protocol operating under a controlled environment can be less restrictive than the requirements of general internet.

LISP-GPE scope of applicability is the same set of use cases covered by[I-D.ietf-lisp-rfc6830bis] for the LISP dataplane protocol. The common property of these use cases is a large set of cooperating entities seeking to communicate over the public Internet or other
large underlay IP infrastructures, while keeping the addressing and topology of the cooperating entities separate from the underlay and Internet topology, routing, and addressing.

LISP-GPE is meant to be deployed in network environments operated by a single operator or adjacent set of cooperating network operators that fits with the definition of controlled environments in [RFC8085].

For the purpose of this document, a traffic-managed controlled environment (TMCE), outlined in [RFC8086], is defined as an IP network that is traffic-engineered and/or otherwise managed (e.g., via use of traffic rate limiters) to avoid congestion. Significant portions of text in this Section are based on [RFC8086].

It is the responsibility of the network operators to ensure that the guidelines/requirements in this section are followed as applicable to their LISP-GPE deployments

4.2. Congestion Control Functionality

LISP-GPE does not natively provide congestion control functionality and relies on the payload protocol traffic for congestion control. As such LISP-GPE MUST be used with congestion controlled traffic or within a network that is traffic managed to avoid congestion (TMCE). An operator of a traffic managed network (TMCE) may avoid congestion by careful provisioning of their networks, rate-limiting of user data traffic and traffic engineering according to path capacity.

Encapsulated payloads may have Explicit Congestion Notification mechanisms that may or may not be mapped to the outer IP header ECN field. Such new encapsulated payloads, when registered with LISP-GPE, MUST be accompanied by a set of guidelines derived from [I-D.ietf-tsvwg-ecn-encap-guidelines] and [RFC6040].

4.3. UDP Checksum

For IP payloads, section 5.3 of [I-D.ietf-lisp-rfc6830bis] specifies how to handle UDP Checksums encouraging implementors to consider UDP checksum usage guidelines in section 3.4 of [RFC8085] when it is desirable to protect UDP and LISP headers against corruption.

In order to provide integrity of LISP-GPE headers, options and payload, for example to avoid mis-delivery of payload to different tenant systems in case of data corruption, outer UDP checksum SHOULD be used with LISP-GPE when transported over IPv4. The UDP checksum provides a statistical guarantee that a payload was not corrupted in transit. These integrity checks are not strong from a coding or
cryptographic perspective and are not designed to detect physical-layer errors or malicious modification of the datagram (see Section 3.4 of [RFC8085]). In deployments where such a risk exists, an operator SHOULD use additional data integrity mechanisms such as offered by IPsec.

An operator MAY choose to disable UDP checksum and use zero checksum if LISP-GPE packet integrity is provided by other data integrity mechanisms such as IPsec or additional checksums or if one of the conditions in Section 4.3.1 a, b, c are met.

By default, UDP checksum MUST be used when LISP-GPE is transported over IPv6. A tunnel endpoint MAY be configured for use with zero UDP checksum if additional requirements in Section 4.3.1 are met.

4.3.1. UDP Zero Checksum Handling with IPv6

When LISP-GPE is used over IPv6, UDP checksum is used to protect IPv6 headers, UDP headers and LISP-GPE headers and payload from potential data corruption. As such by default LISP-GPE MUST use UDP checksum when transported over IPv6. An operator MAY choose to configure to operate with zero UDP checksum if operating in a traffic managed controlled environment as stated in Section 4.1 if one of the following conditions are met:

a. It is known that the packet corruption is exceptionally unlikely (perhaps based on knowledge of equipment types in their underlay network) and the operator is willing to take a risk of undetected packet corruption

b. It is judged through observational measurements (perhaps through historic or current traffic flows that use non zero checksum) that the level of packet corruption is tolerably low and where the operator is willing to take the risk of undetected corruption

c. LISP-GPE payload is carrying applications that are tolerant of misdelivered or corrupted packets (perhaps through higher layer checksum validation and/or reliability through retransmission)

In addition LISP-GPE tunnel implementations using Zero UDP checksum MUST meet the following requirements:

1. Use of UDP checksum over IPv6 MUST be the default configuration for all LISP-GPE tunnels

2. If LISP-GPE is used with zero UDP checksum over IPv6 then such xTR implementation MUST meet all the requirements specified in
section 4 of [RFC6936] and requirements 1 as specified in section 5 of [RFC6936]

3. The ETR that decapsulates the packet SHOULD check the source and destination IPv6 addresses are valid for the LISP-GPE tunnel that is configured to receive Zero UDP checksum and discard other packets for which such check fails.

4. The ITR that encapsulates the packet MAY use different IPv6 source addresses for each LISP-GPE tunnel that uses Zero UDP checksum mode in order to strengthen the decapsulator’s check of the IPv6 source address (i.e., the same IPv6 source address is not to be used with more than one IPv6 destination address, irrespective of whether that destination address is a unicast or multicast address). When this is not possible, it is RECOMMENDED to use each source address for as few LISP-GPE tunnels that use zero UDP checksum as is feasible.

5. Measures SHOULD be taken to prevent LISP-GPE traffic over IPv6 with zero UDP checksum from escaping into the general Internet. Examples of such measures include employing packet filters at the PETR and/or keeping logical or physical separation of LISP network from networks carrying General Internet.

The above requirements do not change either the requirements specified in [RFC2460] as modified by [RFC6935] or the requirements specified in [RFC6936].

The requirement to check the source IPv6 address in addition to the destination IPv6 address, plus the recommendation against reuse of source IPv6 addresses among LISP-GPE tunnels collectively provide some mitigation for the absence of UDP checksum coverage of the IPv6 header. A traffic-managed controlled environment that satisfies at least one of three conditions listed at the beginning of this section provides additional assurance.

4.4. Ethernet Encapsulated Payloads

When a LISP-GPE router performs Ethernet encapsulation, the inner 802.1Q [IEEE.802.1Q_2014] 3-bit priority code point (PCP) field MAY be mapped from the encapsulated frame to the 3-bit Type of Service field in the outer IPv4 header, or in the case of IPv6 the ‘Traffic Class’ field.

When a LISP-GPE router performs Ethernet encapsulation, the inner header 802.1Q [IEEE.802.1Q_2014] VLAN Identifier (VID) MAY be mapped to, or used to determine the LISP Instance IDentifier (IID) field.
5. Backward Compatibility

LISP-GPE uses the same UDP destination port (4341) allocated to LISP.

When encapsulating IP packets to a non LISP-GPE capable router the P-bit MUST be set to 0. That is, the encapsulation format defined in this document MUST NOT be sent to a router that has not indicated that it supports this specification because such a router would ignore the P-bit (as described in [I-D.ietf-lisp-rfc6830bis]) and so would misinterpret the other LISP header fields possibly causing significant errors.

5.1. Detection of ETR Capabilities

The detection of ETR capabilities to support multiple data plane encapsulations and shim headers is out of the scope of this document. Given that the applicability domain of LISP-GPE is a traffic-managed controlled environment, ITR/ETR (xTR) configuration mechanisms may be used for this purpose.

6. IANA Considerations

6.1. LISP-GPE Next Protocol Registry

IANA is requested to set up a registry of LISP-GPE "Next Protocol". These are 8-bit values. Next Protocol values in the table below are defined in this document. New values are assigned under the Specification Required policy [RFC8126]. The protocols that are being assigned values do not themselves need to be IETF standards track protocols.

<table>
<thead>
<tr>
<th>Next Protocol</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00</td>
<td>Reserved</td>
<td>This Document</td>
</tr>
<tr>
<td>0x01</td>
<td>IPv4</td>
<td>This Document</td>
</tr>
<tr>
<td>0x02</td>
<td>IPv6</td>
<td>This Document</td>
</tr>
<tr>
<td>0x03</td>
<td>Ethernet</td>
<td>This Document</td>
</tr>
<tr>
<td>0x04</td>
<td>NSH</td>
<td>This Document</td>
</tr>
<tr>
<td>0x05..0x7F</td>
<td>Unassigned</td>
<td></td>
</tr>
<tr>
<td>0x82..0xFF</td>
<td>Unassigned</td>
<td></td>
</tr>
</tbody>
</table>

7. Security Considerations

LISP-GPE security considerations are similar to the LISP security considerations and mitigation techniques documented in [RFC7835].
LISP-GPE, as many encapsulations that use optional extensions, is subject to on-path adversaries that by manipulating the g-Bit and the packet itself can remove part of the payload. Typical integrity protection mechanisms (such as IPsec) SHOULD be used in combination with LISP-GPE by those protocol extensions that want to protect from on-path attackers.

With LISP-GPE, issues such as data-plane spoofing, flooding, and traffic redirection may depend on the particular protocol payload encapsulated.

8. Acknowledgements and Contributors

A special thank you goes to Dino Farinacci for his guidance and detailed review.

This Working Group (WG) document originated as draft-lewis-lisp-gpe; the following are its coauthors and contributors along with their respective affiliations at the time of WG adoption. The editor of this document would like to thank and recognize them and their contributions. These coauthors and contributors provided invaluable concepts and content for this document’s creation.

- Darrel Lewis, Cisco Systems, Inc.
- Fabio Maino, Cisco Systems, Inc.
- Paul Quinn, Cisco Systems, Inc.
- Michael Smith, Cisco Systems, Inc.
- Navindra Yadav, Cisco Systems, Inc.
- Larry Kreeger
- John Lemon, Broadcom
- Puneet Agarwal, Innovium

9. References

9.1. Normative References

[I-D.ietf-lisp-rfc6830bis]
9.2. Informative References

[I-D.brockners-ippm-ioam-vxlan-gpe]
Brockners, F., Bhandari, S., Govindan, V., Pignataro, C.,
Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Kfir, A.,
Gafni, B., Lapukhov, P., and M. Spiegel, "VXLAN-GPE
Encapsulation for In-situ OAM Data", draft-brockners-ippm-
ioam-vxlan-gpe-02 (work in progress), July 2019.

[I-D.ietf-tsvwg-ecn-encap-guidelines]
Briscoe, B., Kaippallimalil, J., and P. Thaler,
"Guidelines for Adding Congestion Notification to
Protocols that Encapsulate IP", draft-ietf-tnsvwg-ecn-
encap-guidelines-13 (work in progress), May 2019.

[I-D.lemon-vxlan-lisp-gpe-gbp]
Lemon, J., Maino, F., Smith, M., and A. Isaac, "Group
Policy Encoding with VXLAN-GPE and LISP-GPE", draft-lemon-
vxlan-lisp-gpe-gbp-02 (work in progress), April 2019.

(IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460,

[RFC6935] Eubanks, M., Chimento, P., and M. Westerlund, "IPv6 and
UDP Checksums for Tunneled Packets", RFC 6935,
DOI 10.17487/RFC6935, April 2013,
Internet-Draft       LISP Generic Protocol Extension       November 2019


Authors’ Addresses

Fabio Maino (editor)
Cisco Systems
San Jose, CA  95134
USA

Email: fmaino@cisco.com