intarea
Internet-Draft
Intended status: Standards Track
Expires: February 26, 2022

Z. Zhang
R. Bonica
K. Kompella
Juniper Networks
G. Mirsky
ZTE

August 25, 2021

Generic Delivery Functions
draft-zzhang-intarea-generic-delivery-functions-02

Abstract

Some functionalities (e.g., fragmentation/reassembly and Encapsulating Security Payload) provided by IPv6 can be viewed as delivery functions independent of IPv6 or even IP entirely. This document proposes to provide those functionalities at different layers (e.g., MPLS, BIER or even Ethernet) independent of IP.

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 February 10, 2022.

Table of Contents

1. Introduction

Consider an operator providing Ethernet services such as EVPN. The Ethernet frames that a Provider Edge (PE) device receives from a Customer Edge (CE) device may have a larger size than the PE-PE path MTU (PMTU) in the provider network. This could be because

1. the provider network is built upon virtual connections (e.g., pseudowires) provided by another infrastructure provider, or

2. the customer network uses jumbo frames while the provider network does not, or

3. the provider-side overhead for transporting customer packets across the network pushes past the PMTU.

In any case, the provider cannot simply require its customers to change their MTU.

To get those large frames across the provider network, currently, the only workaround is to encapsulate the frames in IP (with or without GRE) and then fragment the IP packets. Even if MPLS is used for service delimiting, IP is used for transportation (MPLS over IP/GRE). This may not be desirable in certain deployment scenarios, where MPLS is the preferred transport or IP encapsulation overhead is deemed excessive.

IPv6 fragmentation and reassembly are based on the IPv6 Fragmentation header below [RFC8200]:

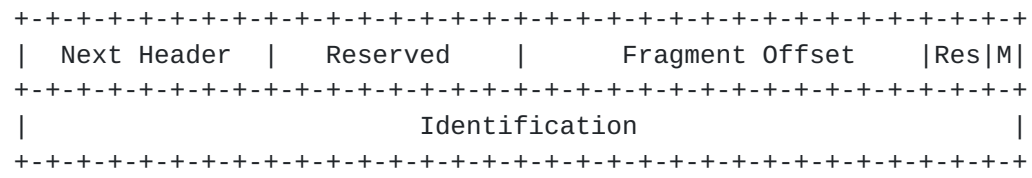


Figure 1: IPv6 Fragmentation Header

This document proposes adapting this header for use in non-IP contexts since the fragmentation/reassembly function is actually independent of IPv6 except for the following aspects:

- The fragment header is identified as such by the "previous" header.
- The "Next Header" value is from the "Internet Protocol Numbers" registry.
- The "Identification" value is unique in the (source, destination) context provided by the IPv6 header.

The "Identification" field, in conjunction with the IPv6 source and destination addresses identifies fragments of the original packet for the purpose of reassembly.

Therefore, the fragmentation/reassembly function can be applied at other layers as long as a) the fragment header is identified as such; and b) the context for packet identification is provided. Examples of such layers include MPLS, BIER, and Ethernet (if IEEE determines it is so desired).

For the same consideration, the IP Encapsulating Security Payload (ESP) [RFC4303] could also be applied at other layers if ESP is desired there. For example, if for whatever reason the Ethernet service provider wants to provide ESP between its PEs, it could do so without requiring IP encapsulation if ESP is applied at non-IP layers.

Similarly, In-Situ OAM (IOAM) functions [I-D.ietf-ippm-ioam-data] can also be applied to many layers.

We refer to these as Generic Delivery Functions (GDFs), which could be achieved at a shim layer between a source and destination delivery points, for example:

- Source and destination IP/Ethernet nodes
- Ingress and egress nodes of MPLS Label Switch Paths (LSPs)

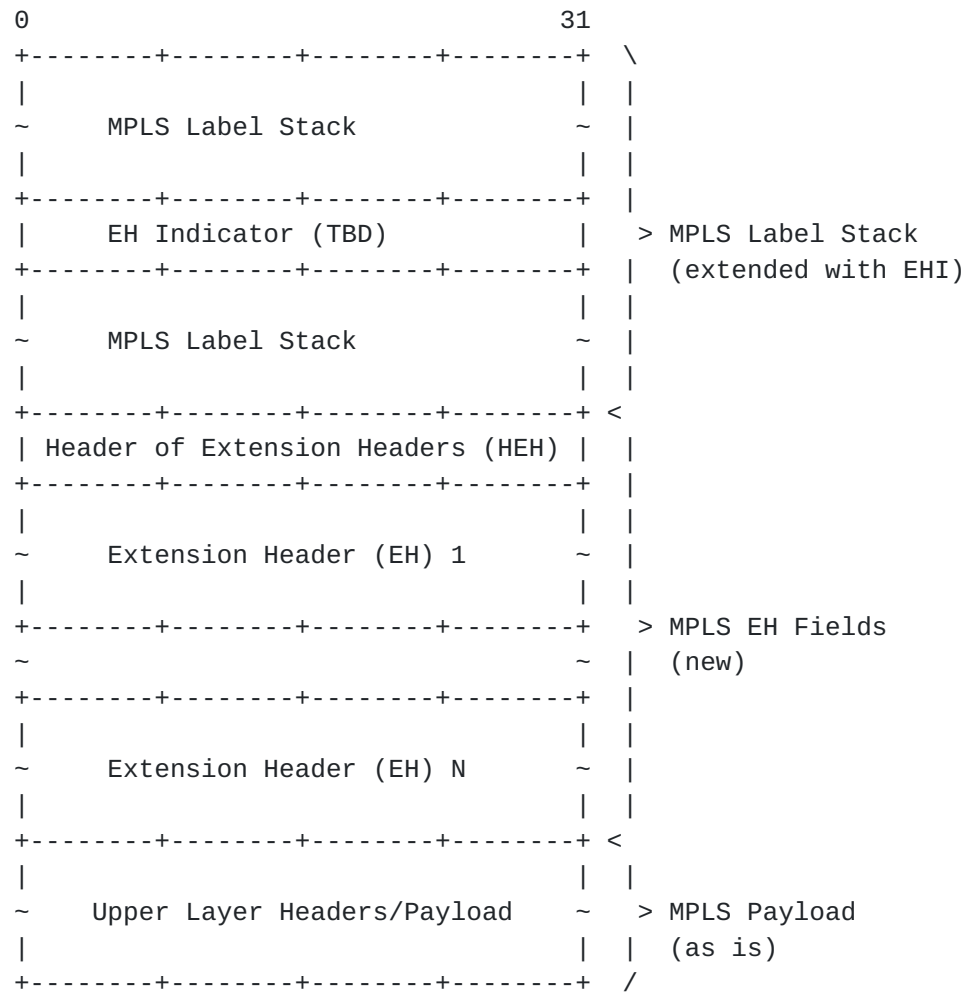
- BIER Forwarding Ingress Routers (BFIRs) and BIER Forwarding Egress Routers (BFERs)

2. Specifications

A Generic Delivery Function, being generic, is likely applicable to IP as well. Therefore, IPv6 Extension Headers (for some GDFs) are directly used at other layers.

2.1. MPLS layer

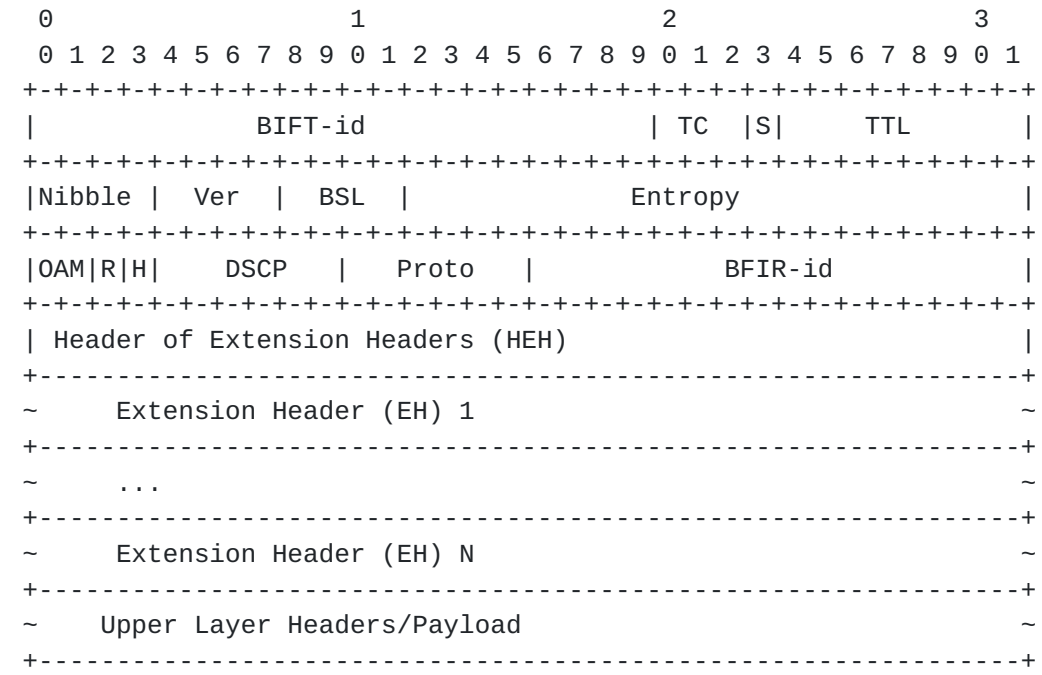
[I-D.song-mpls-extension-header] specifies MPLS Extension Headers encoding. A label entry in the stack indicates the presence of extension headers after the label stack. It starts with a Header of Extension Headers, as depicted in the following excerpt from that specification:



One or more of the EHs in the above can be an IPv6 Extension Header for a GDF.

2.2. BIER layer

For BIER layer, a TBD value for the "proto" field in the outer BIER header indicates that some BIER Extension Headers follow the BIER header, including some IPv6 Extension Headers for GDFs.



R: The "R" flag bit is reserved. It MUST be set to 0 on transmit and ignored on receive.

H: If the "H" flag bit, it indicates the presence of at least one extension header that needs to be processed hop by hop even before a BFER is reached. In this case, the Proto field must be set to the TBD value indicating the presence of extension headers.

2.3. Other layers

Similarly, any layer can have an indication in its packet header that some GDF extension headers follow, including some IPv6 Extension Headers for GDF purpose.

For example, if the outer header is Ethernet (if IEEE would decide to provide the generic delivery functions on top of Ethernet directly), then a new Ethertype would be assigned by IEEE to indicate the presence of GDF extension headers.

2.4. Generic Fragmentation Header (GFH){#gfh}

For generic fragmentation/reassembly functionality, the existing IPv6 Fragment Header needs to be enhanced for MPLS as following:

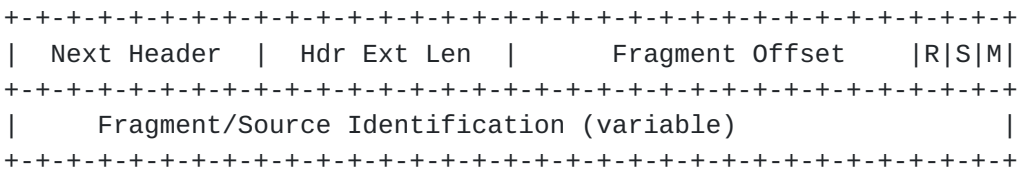


Figure 2: Generic Fragmentation Header

R: The "R" flag bit is reserved. It MUST be set to 0 on transmit and ignored on receive.

S: If the "S" flag bit is clear, the context for the Identification field is provided by the outer header, and only the source-identifying information in the outer header is used.

If the "S" flag bit is set, the variable Identification field encodes both source-identifying information (e.g., the IP address of the node adding the GFH) and an identification number unique within that source. The length of the Fragment header is encoded in the 8-bit "Hdr Ext Len" field (which is a Reserved field in the original IPv6 Fragment Header).

When a GFH is used together with other GDF Headers (GDFH), the GFH SHOULD be the first GDFH.

The above enhancement is not necessary but MAY be used for BIER as well. If the outer header is BIER and the "S" flag bit is clear, the "BFIR-id" field in the BIER header provides the context for the "Identification" field. If the bit is set, then the source information embedded in the source/fragment identification field is used.

3. Security Considerations

To be provided.

4. Acknowledgements

The authors thank Stewart Bryant and Tony Przygienda for their valuable comments and suggestions.

5. References

5.1. Normative References

[I-D.song-mpls-extension-header]
Song, H., Li, Z., Zhou, T., Andersson, L., and Z. Zhang, "MPLS Extension Header", draft-song-mpls-extension-header-05 (work in progress), July 2021.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <https://www.rfc-editor.org/info/rfc5308>.

[RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <https://www.rfc-editor.org/info/rfc7684>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", RFC 8362, DOI 10.17487/RFC8362, April 2018, <https://www.rfc-editor.org/info/rfc8362>.

5.2. Informative References

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-14 (work in progress), June 2021.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, DOI 10.17487/RFC4303, December 2005, <https://www.rfc-editor.org/info/rfc4303>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

Ron Bonica
Juniper Networks

Email: rbonica@juniper.net

Kireeti Kompella
Juniper Networks

Email: kireeti@juniper.net

Gregory Mirsky
ZTE

Email: gregory.mirsky@ztetx.com