idr
Internet-Draft
Intended status: Standards Track
Expires: 10 August 2023

Z. Zhang
Juniper Networks
S. Hares

6 February 2023

MPLS Label Stacks in Tunnel Encapsulation Attribute
draft-zzhang-idr-tunnel-encapsulation-label-stack-02

Abstract

RFC 9012 defines an MPLS Label Stack sub-TLV for Tunnel Encapsulation Attribute, and specifies that it is to be pushed BEFORE other labels. This document clarifies the use case for that, defines a new Tunnel Label Stack sub-TLV for a label stack to be pushed AFTER other labels (e.g., the label embedded in the NLRI for a labeled address family, and/or the stack in an MPLS Label Stack sub-TLV), and defines two new Segment sub-TLVs to encode a segment list in a compact format.

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 10 August 2023.

Copyright Notice

Table of Contents

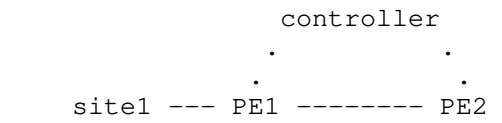
1. Traffic Steering after Tunnel Endpoint

[RFC9012] defines an MPLS Label Stack sub-TLV for Tunnel Encapsulation Attribute and specifies that:

"If a packet is to be sent through the tunnel identified in a particular TLV, and if that TLV contains an MPLS Label Stack sub-TLV, then the label stack appearing in the sub-TLV MUST be pushed onto the packet before any other labels are pushed onto the packet."

Specifically, the label stack in the sub-TLV is to be pushed BEFORE any other labels are pushed. This may sound counter-intuitive, in that if a label stack (e.g. for Segment Routing) is to be used to steer traffic to the tunnel endpoint, the stack should be pushed AFTER other labels (e.g. the label embedded in the NLRI).

This document clarifies that it is NOT for steering traffic to but for steering AFTER the tunnel endpoint. Consider the following use case:



Two sites are connected to two PEs respectively, and traffic steering is desired within each site not just among the PEs. While PE2 could push the label stack used for steering within site2, there may be situations where PE2 is not a device capable of pushing a large label stack so PE1 is tasked with pushing the label stack that is used after the tunnel end point PE2, within site2.

2. Traffic Steering to the Tunnel Endpoint

Notice that in the above example, it may be desired that PE2 or the controller wants PE1 to send service traffic to PE2 via a specific path through the underlay network. The path may be an Segment Routing path either as a pre-installed SR Policy tunnel in the Tunnel Encapsulation Attribute (TEA), or as a label stack encoded in an MPLS tunnel in the TEA of the service routes that PE1 receives. There are different use cases for the two approaches - many TEAs could reference a common SR Policy that has been pre-installed via means in [I-D.ietf-idr-segment-routing-te-policy], or a TEA can simply specify an ad-hoc label statck without having to have an SR policy pre-installed.

In this case, PE1 needs to impose the label stack AFTER it imposes other labels like service labels or the labels for traffic steering at site2 after the traffic arrives at PE2. Obviously, a new sub-TLV is needed to encode the label stack for steering traffic to the tunnel endpoint, as the existing MPLS Label Stack sub-TLV is for steering after traffic reaches the tunnel endpoint.

Notice that, if the routes are advertised by PE2 and received by many other PEs, the path identified in the TEA needs to be a partial path that are closer to PE2 (so that all ingress PEs can still use that path). Otherwise, the more appropriate use case is when the routes are advertised from the controller - whether the routes are for unicast or for programming a multicast replication branch on a router where the downstream node for that branch needs to be reached via a specific path [I-D.ietf-pim-sr-p2mp-policy].

2.1. Tunnel Label Stack sub-TLV

This document defines a new Tunnel Label Stack sub-TLV for that purpose. It has exactly the same syntax as the existing MPLS Label Stack sub-TLV. Section 3.6 of [RFC9012] applies to this new sub-TLV, with the following differences:

* A new tunnel type will be allocated by IANA

* The label stack MUST be imposed AFTER other labels are pushed.

Both the existing MPLS Label Stack sub-TLV and the new Tunnel Label Stack sub-TLV MAY be present in a tunnel TLV.

2.2. SR Policy Tunnel

In [I-D.ietf-idr-segment-routing-te-policy], an SR Policy tunnel type is specified to be used in a TEA attached to an NLRI of SR Policy SAFI.

The SR Policy SAFI is used to install an SR Policy to a node, specifying all applicable properties of that policy like policy name, candidate path, segment list, etc.. After the policy is installed, it can be used to steer traffic into the tunnel.

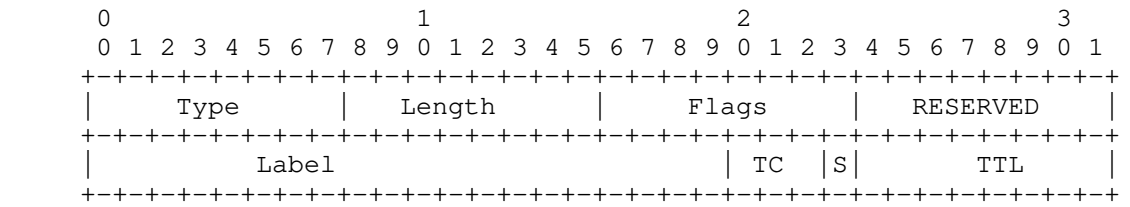
For the use case mentioned earlier, where a tunnel in a TEA for a service route (that is not of the SR Policy SAFI) needs to follow a particular SR path defined in an SR policy that has been pre-installed via an SR Policy SAFI NLRI, the service route's TEA can include an SR Policy tunnel, which MUST only include a policy name sub-TLV, and a receiving router uses the policy name to resolve a pre-installed SR policy.

In other words, the SR Policy tunnel in [I-D.ietf-idr-segment-routing-te-policy] is used to install the policy, while the SR Policy tunnel in this document is for referencing a pre-installed policy. In this version of the document, the same SR Policy tunnel type is used (though only the policy name and nothing else is included), but we could specify a new tunnel type instead depending on WG consensus.

3. Compact Segment List

Section 2.4.4 of [I-D.ietf-idr-segment-routing-te-policy] specifies a Segment List sub-TLV that is optional in a tunnel TLV. It encodes a segment list in an SR Policy tunnel, containing zero or more Segment sub-TLVs.

Each Segment sub-TLV specifies a segment of various types defined in Section 4 of [RFC9256]. For example, if a segment list is a 10-label stack, then the Segment List sub-TLV for it has 10 sub-TLVs, each being a Type A Segment as defined in 2.4.4.2.1 of [I-D.ietf-idr-segment-routing-te-policy]:



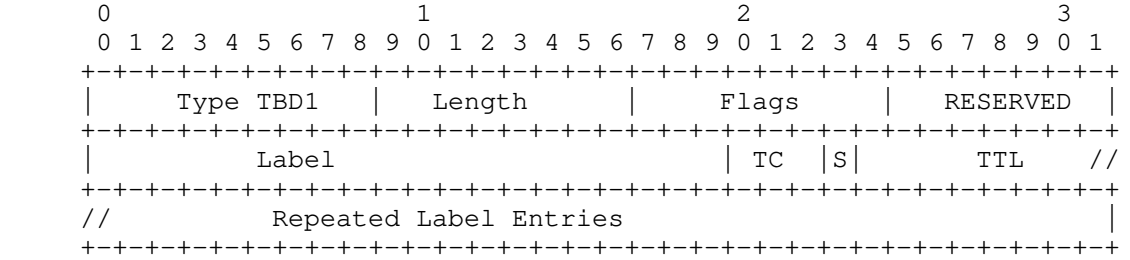
It is clear that this is not efficient on-the-wire encoding format, and it involves additional encoding/decoding processing.

To address this inefficiency, this document specifies two new types of Segment sub-TLVs, each encoding a label stack or an SRv6 SID list respectively. A new segment type may be added in a future revision to encode a compressed SRv6 SID list.

Note that, while each new type is called a Segment sub-TLV in a Segment List sub-TLV, it actually encodes a segment list (a label stack or an SRv6 SID list). A Segment List sub-TLV MAY have a mixed Segment sub-TLVs of any types, e.g., a Type A segment that encodes one label and another new segment type that encodes a label stack. The actual segment list is a concatenation of all the labels in this example.

3.1. Segment Type L

The Type L Segment Sub-TLV encodes multiple SR-MPLS SIDs. The format is as follows and is used to encode MPLS Label fields as specified in [RFC3032] [RFC5462]:



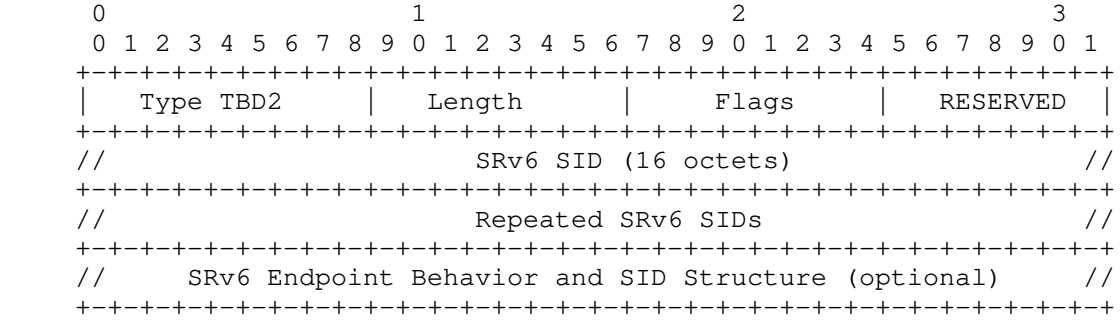
The Type TBD1 is to be assigned by IANA from the "SR Policy Segment List Sub-TLVs" under the "BGP Tunnel Encapsulation" registry.

The Length value is (4 * number of labels + 2).

Other fields are as defined in 2.4.4.2.1 of [I-D.ietf-idr-segment-routing-te-policy].

3.2. Segment Type M

The Type M Segment Sub-TLV encodes multiple SRv6 SIDs with optional SRv6 Endpoint Behavior and SID Structure:



The Type TBD2 is to be assigned by IANA from the "SR Policy Segment List Sub-TLVs" under the "BGP Tunnel Encapsulation" registry.

The Length value is (16 * number of SIDs + 2) when SRv6 Endpoint Behavior and SID Structure is not present. If it is present, the Length value is increased by 8.

Other fields are as defined in 2.4.4.2.2 of [I-D.ietf-idr-segment-routing-te-policy].

4. Security Considerations

This document does not introduce any new security issues besides what is already discussed in RFC9012 and [I-D.ietf-idr-segment-routing-te-policy].

5. IANA Assignments

IANA is requested to assign a new sub-TLV type for "Tunnel Label Stack" from "BGP Tunnel Encapsulation Attribute Sub-TLVs" registry, in the 0~127 range.

IANA is requested to assign two new sub-TLV types from the "SR Policy Segment List Sub-TLVs" under the "BGP Tunnel Encapsulation" registry, for Type L and Type M segments respectively.

6. Acknowledgements

The authors thank Eric Rosen, John Scudder and Robert Raszuk for their valuable comments and suggestions.

7. References

7.1. Normative References

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

[I-D.ietf-idr-segment-routing-te-policy] Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-segment-routing-te-policy-20, 27 July 2022, <https://www.ietf.org/archive/id/draft-ietf-idr-segment-routing-te-policy-20.txt>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

7.2. Informative References

[I-D.ietf-pim-sr-p2mp-policy] Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. J. Zhang, "Segment Routing Point-to-Multipoint Policy", Work in Progress, Internet-Draft, draft-ietf-pim-sr-p2mp-policy-05, 2 July 2022, <https://www.ietf.org/archive/id/draft-ietf-pim-sr-p2mp-policy-05.txt>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net

Susan Hares
Email: skh@ndzh.com