Multicast/BIER As A Service
draft-ietf-bier-multicast-as-a-service-01

Abstract

This document describes a framework for providing multicast as a service via Bit Index Explicit Replication (BIER) [RFC7279], and specifies a few enhancements to [draft-ietf-bier-idr-extensions] [RFC8279] [RFC8401] [RFC8444] to enable multicast/BIER as a service.

Requirements Language

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.

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 January 12, 2023.
Copyright Notice

Copyright (c) 2022 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.  Terminologies ................................................. 3
   1.2.  A CDN of A Single Provider .................................. 4
       1.2.1.  IGP/BGP Interworking ................................. 5
   1.3.  A CDN That Involves Another Provider ....................... 6
       1.3.1.  Providing Independent BAAS To Multiple Customers .. 6
       1.3.2.  BIER VPN ................................................ 7
       1.3.3.  Control and Accounting ............................... 9
   1.4.  Sets and Segmentation ........................................ 9
       1.4.1.  Multiple Sets ......................................... 9
       1.4.2.  Segmentation ........................................... 10
2.  Specifications for Enhancements to BIER Signaling with BGP/IGP ........................................... 10
   2.1.  BGP Procedures .............................................. 10
   2.2.  ISIS/OSPF Procedures ....................................... 11
3.  IANA Considerations ............................................. 12
4.  Security Considerations .......................................... 12
5.  Acknowledgements ................................................ 13
6.  References ........................................................ 13
   6.1.  Normative References ..................................... 13
   6.2.  Informative References ................................... 14
Authors’ Addresses .................................................. 14

1.  Introduction

Currently multicast is primarily used in the following scenarios:

- Enterprise Applications. For example, large scale financial data publishing.
- Provider/underlay tunnels for MVPN and for EVPN BUM.
o Real-time IPTV offered by a service provider to its customers.

Besides the above, large scale multicast services, especially transit multicast transport provided by large Internet Service Providers is virtually non-existent. This is mainly because of the following chicken and egg dilemma:

o Traditional multicast technologies are complicated and lack scalability. The revenue that multicast services bring in cannot offset the Capex and Opex that an operator has to invest, so provider networks typically do not enable multicast even though the deployed equipment does support multicast.

o As a result, Content Providers cannot take advantage of multicast and instead use less efficient methods like Ingress Replication, Peer2Peer, or multicast at application layer.

A recent multicast technology breakthrough, BIER, provides a simple and scalable solution for large scale multicast deployment, independent of number of multicast flows. In the meantime, large scale distribution of ultra high definition video content has become more and more popular and important. Service providers simply cannot keep on increasing their network capacity even if they could shift cost to Content Providers. With these developments, service providers now have both the need and means to provide scalable multicast service, potentially across multiple providers.

This document describes a framework for Multicast As A Service (MAAS) enabled by BIER. We use Content Delivery Network (CDN) as example, though it applies to any large scale multicast delivery service.

1.1. Terminologies

Readers are assumed to be familiar with multicast, BIER, BGP and ISIS/OSPF concepts and procedures. Some terminologies are listed here for convenience.

o BFR: BIER Forwarding Router.

o BFIR: BIER Forwarding Ingress Router.

o BFER: BIER Forwarding Egress Router.

o EBFR: Edge BFR. Including BFIR and BFER.

o BSL: BitStrengLength. Number of bits in the BitString of a BIER header.
1.2. A CDN of A Single Provider

To make it easier to understand, we first consider a simple example: a CDN owned by a single operator, which could be a Content Provider itself. The network spans multiple ASes as shown in the following figure:

```
++++                                       ++++
|EBFR11+   +EBFR12                     EBFR21+   +EBFR22
 +     +                               +     +
 +     +                               +     +
 + AS100   + AS200    +                 +     +
 +     +                               +     +
 +     +                               +     +
 + ASBR132   +++++++                     ASBR231+   +EBFR23
EBFR13   +   +   +   /                 +++
 +     + ASBR312    ASBR321
++++ ASBR131   +       \
|   + AS300    +
| ASBR311 (BFR) +
| ASBR341    ASBR351 (BFR)
/     +       +
++++ /   +++++++   \   ++++
|EBFR43 Torrent ASBR431 ASBR531 EBFR53
 |      +       +
 | AS400   + AS500    +
 |     +       +
EBFR41 EBFR42 EBFR51 EBFR52
++++     +     +     +
```

The CDN uses BIER for multicast transport and Edge BIER Forwarding Routers (EBFRs) are located throughout the network. Some of them are connected towards multicast content sources and are referred to as BIER Forwarding Ingress Routers (BFIRs) in BIER architecture. Most of them are connected towards multicast content receivers and are referred to as BIER Forwarding Egress Routers (BFERs). Notice that between content sources and BFIRs there may be Protocol Independent Multicast (PIM) in use, while between content receivers and BFERs there may be PIM and/or IGMP in use.

At the initial deployment stage, there might be only a few transit BIER Forwarding Routers (BFRs) at strategic points in the network (e.g. ASBR311 and ASBR351). BGP sessions are established among the EBFRs and BFRs, and BGP extensions as defined in
[I-D.ietf-bier-idr-extensions] are used to signal BIER information. All these are in a single BIER sub-domain.

In the example of initial stage with only ASBR311 and ASBR351 as BFRs, multicast traffic arriving at EBFR11 will be imposed with a BIER header and replicated to EBFR12/EBFR13/ASBR311 over tunnels. ASBR311 will further replicate traffic to ASBR351/EBFR41/EBFR42/EBFR43/EBFR21/EBFR22/EBFR23 over tunnels, and ASBR351 will further replicate traffic to EBFR51/EBFR52/EBFR53 over tunnels.

The BGP signaling and a necessary enhancement can be explained using the following example. EBFR43 advertises its BIER prefix (a loopback address) as /32 IPv4 or /128 IPv6 prefix in BGP with a BIER Path Attribute (BPA) [RFC8279] [I-D.ietf-bier-idr-extensions]. ASBR431 receives it and re-advertises it (with BGP Next Hop changed to itself) but does not do anything wrt BIER because it does not support BIER. Same happens on ASBR341. When ASBR311 and ASBR351 receive it from ASBR341, they create a BIFT entry corresponding to EBFR43’s BFR-ID. The entry causes a BIER packet with corresponding bit set in its BitString to be tunneled to EBFR43. This cannot be based on BGP Next Hop in the advertisement because the BGP Next Hop is ASBR341. When eventually EBFR11 receives the re-advertised route, it creates a BIFT entry that causes corresponding packets to be tunneled to ASBR311 (but not to EBFR43 directly). Now it is clear that this cannot be based on either the BIER prefix itself or the BGP Next Hop. The solution is that the originating EBFR attaches a Tunnel Encap Attribute (TEA) [RFC9012] with the tunnel destination set to itself, and whenever a BFR re-advertises the route it changes the tunnel destination to itself. When a BFR creates the BIFT entry, it uses the Tunnel Egress Endpoint in the TEA to find out where to tunnel packets.

Over time, more routers in network may be upgraded to support BIER and become a BFR. For example, once ASBR431 is upgraded to a BFR, ASBR311 no longer needs to tunnel traffic to EBFR41/EBFR42/EBFR43 but only need to tunnel one copy to ASBR431, who will then replicate to EBFR41/EBFR42/EBFR43.

1.2.1. IGP/BGP Interworking

Additionally, if enough routers in an AS (or just one of its IGP areas) can be upgraded to run BIER, then hop-by-hop BIER forwarding can be utilized there, using IGP extensions for BIER signaling [RFC8401] [RFC8444].

Notice that even with this there is still only one BIER sub-domain, with mixed IGP and BGP signaling for BIER. To redistribute BIER
information between IGP and BGP, procedures specified in
[I-D.ietf-bier-prefix-redistribute] and detailed in Section 2.2 are
followed.

1.3. A CDN That Involves Another Provider

In the above example, the CDN is providing multicast transport
service, with simplicity and scalability provided by BIER (the per-
flow state is confined to the edges). Now let us go one step further
and consider that AS300 belongs to a different Internet Service
Provider. Now the ISP is providing BIER As A Service (BAAS) to the
CDN, by being part of the CDN’s BIER sub-domain. Notice that, not
only does the ISP not have per-tree state (it does not have EBFRs),
but also its BFRs do not need BFR-ID assigned. The ISP does need to
learn about all the EBFRs and their corresponding BFR-IDs (through
signaling).

1.3.1. Providing Independent BAAS To Multiple Customers

Now consider that the ISP also provides BAAS for another CDN. Each
of the two CDNs has its own BIER domain, with its own BFR-ID or even
sub-domain ID assignment that could conflict between the two CDNs.
For example, both have BFR-ID 100 and sub-domain ID 0 assigned but
they are totally independent of each other. For an BFR in the ISP to
support this, with BGP signaling it needs to advertise its own BFR
prefix multiple times, each time with a different RD that is mapped
to the corresponding CDN. A new SAFI BIER (to be allocated by IANA)
is used.

In the above example, there are two paths between AS100 and AS300.
It is possible that while ASBR311 is the BFR, ASBR312 is the unicast
best path into AS300 and beyond from AS100. Advertising BIER
prefixes using a different SAFI with a RD also has the side benefit
of allowing incongruent topologies for unicast and BIER.

In the existing BIER architecture and IGP extensions for BIER a sub-
domain is tied to a single topology (either the one and only topology
if Multi-topology ISIS/OSPF is not used, or a topology as defined in
Multi-topology ISIS/OSPF). In the BIER sub-TLV that ISIS/OSPF
attaches to a BIER prefix, a Sub-domain-ID value can only appear once
for a particular topology. In this document, a BFR in the BAAS
provider may belong to different and independent BIER domains, and
the same sub-domain ID needs to be signaled multiple times, once for
each BIER domain (notice that the same sub-domain-ID actually
identifies different sub-domains in different BIER domains, so this
does not really change the architectural requirement that a sub-
domain is tied to a single topology). To do so, a new "BIER Domain"
sub-TLV is introduced, and its value field includes a RD (as in the
BGP signaling) and a BIER sub-sub-TLV that is the same as currently specified in ISIS/OSPF extensions for BIER.

This works very well because of the flexible BIER architecture - a BIER packet is forwarded based on a Bit Index Forwarding Table (BIFT) that is determined by a 20-bit BIFT ID in front of the BIER header, and each (subdomain, BSL, set) tuple has its own BIFT. Traditionally, a subdomain is identified by a sub-domain ID but in this document a subdomain is now identified by a (RD, sub-domain ID) tuple in the control plane.

With this, the scaling aspect on a BFR comes to how many BAAS customer the provider needs to support. For example, if it needs to support 16 BAAS customers, one BSL, and four sets (Section 1.4.1) for each customer, then the provider needs to support 64 BIFTs (16 x 1 x 4). If the BSL is 256, then each BIFT has 256 entries in it and the total number of BIFT entries (routes) is 4k (256 x 64). Notice that this 4k number is not related to the number of customers' multicast flows, but only related to the number of customers and number of customer EBFRs. The number of customers with their own independent BIER domains are likely not very large initially, but if multicast as a service gets more widely used, the protocol and procedures defined in this document can scale up to the extent of how many BIFTs (and BIFT entries) a BFR can support. Since there is no real difference between a BIFT entry and a unicast RIB/FIB entry, as long as the scaling requirements are adequately considered in the BIER forwarding plane implementation (e.g., enough memory is allocated for the BIFTs), scaling will not become a bottleneck.

Building/updating the BIFTs is the same as in the base BIER architecture, except that in the control plane a subdomain is identified by a (RD, sub-domain ID) tuple instead of just a sub-domain ID. This is transparent to the forwarding plane - a BIFT is always identified by an opaque 20-bit opaque number. This opaque number is either a label for MPLS encapsulation or an opaque number for non-MPLS encapsulation, and the optional static encoding as specified in [I-D.ietf-bier-non-mpls-bift-encoding] cannot be used.

1.3.2. BIER VPN

In the previous section, customer BIER prefixes (with non-zero BFR-IDs) are advertised into the provider network with BIER information. Notice that it is not a VPN scenario - even though the customers and the provider may have independent and overlapped BIER sub-domain and BFR-IDs, the customer networks and the provider network are in the same address space.
While the same mechanism can be used for VPN, this section looks at VPN scenario at a different angle and considers more aspects. It will become clear that these apply to non-VPN scenario as well.

1.3.2.1. Signaling in/over Provider Underlay

Advertising non-VPN customer BIER prefixes into provider network as described in Section 1.3.1 may be considered acceptable (this is just like advertising internet routes), but advertising VPN customer BIER prefixes into provider network is not aligned with VPN architecture [RFC4364].

Notice that while BIER signaling attaches BIER information to BIER prefixes that are used to find routing underlay paths to BFERs, it is the BFER-IDs that are used in BIFTs. In other words, the key is to be able to find paths to BFERs that are identified as BFR-IDs (or bit positions in BIER packets’ BitString to be more precise). Since customers BFERs will be reached via PEs, as long as P routers know which customer BFR-IDs are associated with (i.e., reachable via) which PEs, they can calculate the BIFTs for the customer BIER domains without knowing customer BIER prefixes. This association information can be advertised by PEs via Proxy range sub-TLV [I-D.ietf-bier-prefix-redistribute].

The signaling discussed above is in the provider underlay, with customer BFR-IDs advertised with PE BIER prefixes. In parallel, with overlay signaling via MPLS-labeled VPN address SAFI [RFC4364] among PEs, BIER Path Attribute [I-D.ietf-bier-idr-extensions] carries customer BIER information (including customer BIER prefixes) from site to site, with redistribution [I-D.ietf-bier-prefix-redistribute] between IGP and BGP at PEs.

The underlay/overlay signaling is illustrated in the following figure. Customer sites run customer IGP/BGP towards the VRFs on PEs, with customer BIER information advertised with customer BIER prefixes. Overlay signaling then happens among PEs, with the complete BIER information carried in BIER Path Attribute attached to customer BIER prefixes (that are advertised as normal VPN routes), and redistribution happens at the PEs. In the mean time, underlay signaling among PE/P routers advertises customer BFR-ID and sub-domain information via BIER proxy range sub-TLVs.

--- overlay BGP ---
  /                   \
site 1 ---- vrf|PE1      P      PE2|vrf ---- site 2
  IGP/BGP     \    / \
  \     /  IGP/BGP
  ----  -----  ----
underlay IGP/BGP
1.3.2.2. VPN vs. Non-VPN

As alluded to earlier, the VPN procedures (in particular using BIER Proxy Range sub-TLV) can be used for non-VPN scenarios as well, and the overlay signaling for the non-VPN case can be via BGP-LU [RFC3107] with BIER Path Attribute attached.

1.3.3. Control and Accounting

With BGP based signaling, internal routers of a BAAS provider does not need explicit configuration for the BIER transport services that it support. In the above example, the ASBRs (ASBR311, ASBR312, ASBR321, ASBR341, ASBR351) in AS300 only need to have BGP policy configured to allow certain received BIER prefix advertisements to trigger necessary BIER state and additional signaling of their own. For example, when ASBR351 receives the BIER prefix advertisement, if its local configuration allows it may create corresponding BIFTs and BIFT entries, and additionally originates or updates its own BIER prefix advertisement. An internal BFR inside AS300, upon receiving the BGP advertisements, may or may not need to go through the same policy check again (based on the providers operation model).

When the ASBRs (re-)advertise BIER prefixes toward their external peers, they could enable statistics counters for the corresponding BIER labels so that they can count incoming BIER packets from external peers specifically for this BAAS. Similarly, the ASBRs can enable statistics counters for BIER labels they receive from external peers, so that they can count outgoing BIER packets delivered to the external peers. These incoming and outgoing counters can be used for accounting and billing purposes.

1.4. Sets and Segmentation

The number of EBFRs could very well be larger than the BSL. There are two ways to handle that - multiple sets or segmentation.

1.4.1. Multiple Sets

With this method the set of EBFRs are grouped into multiple sets, and the number of EBFRs in a set is smaller than the BSL. A BFIR may need to send multiple copies of a multicast packet to reach all BFERs, one copy for each set that covers one or more expecting BFERs. A separate BIFT is needed for each set (because the same bit in the BitString of packets for different sets maps to different BFERs).

This not only leads to multiple copies to be sent over the same link, but also requires additional BIFTs. In the earlier example, 64 BIFTs are needed for 16 BAAS customers because each customer needs 4 BIFTs for the multiple sets.
1.4.2. Segmentation

With this method, a BIER network is segmented into multiple regions, each with its own BIER sub-domain. In the earlier example, each AS could be an independent sub-domain. A BIER packet from EBFR11 will be decapsulated by the segmentation border router ASBR311, and then sent into next sub-domain in AS300 with a new BIER header. The segmentation [RFC7524] involves Multicast Flow Overlay [RFC8279] [RFC8556] so that the segmentation border routers know what BitString to use when sending onto the next segment. The advantage of segmentation is that only a single copy needs to be sent, and the number of BIFTs is also reduced on all BFRs. The disadvantage is that the segmentation points need to run multicast flow overlay protocol and maintain related state in control plane and data plane.

A deployment may start without the need for either multiple sets or segmentation when the number of EBFRs is small. When the number of EBFRs grows, segmentation can be introduced incrementally. A new BFR can be added as, or an existing BFR could be converted to, a segmentation point, splitting the original sub-domain into two independent sub-domains. The segmentation point does not re-advertise BIER information from one sub-domain to another. Other BFRs/EBFRs do not need any configuration changes except to make sure that all BIER information exchange is restricted to a single sub-domain (for example, two BFRs were BGP peers before and were exchanging BIER information but now they belong to two sub-domains and only exchange BIER information with the segmentation point and other BFRs in the same sub-domain).

In the earlier example of a CDN of a single provider, using segmentation may be acceptable, even though the overlay state needs to be kept by the segmentation points. A BAAS provider may need to carefully consider if it wants to keep a customer's overlay state on those segmentation points. On the other hand, the provider may consider hosting per-customer segmentation points. For example, tethering small or virtual BFRs to an ASBR and have those BFRs be the segmentation points [I-D.ietf-bier-tether].

2. Specifications for Enhancements to BIER Signaling with BGP/IGP

2.1. BGP Procedures

When an EBFR advertises a BIER prefix with a BIER Path Attribute (BPA), it SHOULD attach a Tunnel Encap Attribute (TEA) with the tunnel destination set to itself.

A BFR receiving the advertisement MUST use the tunnel destination in the TEA to determine where to forward a BIER packet whose BitString
has a set bit corresponding to the BIER prefix, unless the TEA does not exist, in which case the BIER prefix itself is used for the determination. When the BFR re-advertises the BIER prefixes, it MUST change the tunnel destination in the TEA to itself, or add a TEA with the tunnel destination set to itself if there was no TEA in the received advertisement.

The TEA SHOULD have a Protocol Sub-TLV with protocol type BIER (0xAB37).

A transit BFR that is allowed (by provisioning or based on policy) to participate in a BIER sub-domain MUST advertise its own BIER prefix with a BPA. The BFR-id in the BPA SHOULD be 0. Depending on the operational model of the operator, the advertisement MAY be based on received BIER prefixes (subject to certain BGP policy verification), or MAY do so only with explicit configuration.

If a provider provides independent BAAS services to multiple customers, when its BFR receives BIER prefixes from a customer it MUST re-advertise with a new BIER SAFI. For simplicity, all BFRs of the provider use the same RD that is specifically assigned for the customer. When a BFR re-advertises BIER prefixes to a customer, it MUST re-advertise with SAFI 1 or 2.

If multiple providers together provide BAAS to a customer, then the two providers may assign the same RD for the customer or do RD rewriting when re-advertising BIER prefixes from one provider to another.

2.2. ISIS/OSPF Procedures

This document defines a new BIER Domain Sub-TLV of ISIS TLVs 135, 235, 236, and 237. The sub-TLV type is to be allocated.

This document also defines a new BIER Domain Sub-TLV of OSPF Extended Prefix TLV. The sub-TLV type is to be allocated.

The value part of the BIER Domain Sub-TLV includes a 64-bit Route Distinguisher followed by one or more BIER Info Sub-TLV (as defined in [RFC8401] and [RFC8444] respectively) as its sub-sub-TLVs.

When a BFR redistribute a BIER prefix from BGP into ISIS/OSPF, if the BGP advertisement is of BIER SAFI, a BIER Domain sub-TLV is attached, with the RD part of the sub-TLV copied from the BGP advertisement. For each BIER TLV in the BPA, a BIER Info sub-sub-TLV is added in the BIER Domain sub-TLV, with the subdomain-id and BFR-id copied from the corresponding BIER TLV in the BPA, and the Encapsulation sub-sub-sub-TLV omitted because it is not needed.
If the BGP advertisement is of SAFI 1 or 2, BIER Info Sub-TLVs are constructed as above directly, without using a BIER Domain sub-TLV.

When a BFR redistribute a BIER prefix from ISIS/OSPF into BGP, if there is a BIER Domain sub-TLV in the corresponding ISIS LSP or OSPF LSA, the BGP advertisement is of BIER SAFI and the RD part of the NLRI is set to the RD from the BIER Domain sub-TLV. For each BIER Info sub-sub-TLV in the BIER Domain sub-TLV, a BIER TLV is included in the BPA, with the subdomain-id and BFR-id copied from the corresponding BIER Info sub-sub-TLV. The MPLS Encapsulation sub-TLV is omitted. The tunnel destination in the TEA is set to the BFR’s BIER prefix.

If there is no BIER Domain sub-TLV in the corresponding ISIS LSP or OSPF LSA for the BIER Prefix, the BGP advertisement is of SAFI 1 or 2, and the BPA is constructed similar to the above (the only difference is that in this case BIER Info sub-TLVs are not part of a BIER Domain sub-TLV).

3. IANA Considerations

This document requests the following IANA assignments:

- A sub-TLV type for BIER Domain Sub-TLV from ISIS "Sub-TLVs for TLVs 135, 235, 236, and 237" registry.
- A sub-TLV type for BIER Domain Sub-TLV from OSPFv2 Extended Prefix Sub-TLV registry.
- A BIER SAFI from Subsequent Address Family Identifiers (SAFI) registry.

4. Security Considerations

There are no security concerns wrt exchange of BIER information besides what have been discussed in [I-D.ietf-bier-idr-extensions] and [RFC8401] [RFC8444].

The tunnels between BFRs that are not directly connected are ideally auto-configured to reduce provisioning burdens. Given that they may span multiple ASes and MPLS may not always be available, BIER over UDP/GRE/IPv4/IPv6 becomes very convenient, though that has the same security concerns well discussed in "Security Considerations" of [RFC4023] and [RFC7510].

As one mitigation when the tunnel is not secured, a BFR MAY use source address filtering based on pre-provisioned or dynamically learned allowable addresses. With dynamic learning, if a BFR
receives a BIER prefix with a BPA and a TEA (see Section 2.1), it sets up a forwarding filter to allow IP/GRE/UDP tunneling from the address encoded in the "Tunnel Egress Endpoint" sub-TLV of Tunnel TLVs in the TEA. While that is the address for this BFR to tunnel traffic to, this BFR will also likely receive tunneled traffic from that address.

5. Acknowledgements

The authors thank Lenny Giuliano and Antoni Przygenda for their review and suggestions.

6. References

6.1. Normative References

[I-D.ietf-bier-idr-extensions]

[I-D.ietf-bier-prefix-redistribute]


6.2. Informative References

[I-D.ietf-bier-non-mpls-bift-encoding]

[I-D.ietf-bier-tether]


Authors’ Addresses

Zhaohui Zhang
Juniper Networks
EMail: zzhang@juniper.net

Eric Rosen
Individual
EMail: erosen52@gmail.com
IS-IS extensions for BIER-TE (Tree Engineering for Bit Index Explicit Replication) with MPLS and non-MPLS Encapsulation
draft-zwx-bier-te-isis-extensions-01

Abstract

This document describes the IS-IS protocol extension that is required for BIER-TE with MPLS and non-MPLS 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 25 January 2023.

Copyright Notice

Copyright (c) 2022 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.
1.  Introduction

[I-D.ietf-bier-te-arch] describes per-packet stateless strict and loose path steered replication and forwarding for "Bit Index Explicit Replication" (BIER, [RFC8279]) packets. It is called BIER Tree Engineering (BIER-TE) and is intended to be used as the path steering mechanism for Traffic Engineering with BIER.

BIER-TE introduces a new semantic for "bit positions" (BP) of link. [I-D.ietf-bier-te-isis] describes IS-IS extensions for distributing BitPositions configured on the links in "Bit Index Explicit Replication Traffic Engineering" (BIER-TE) domain.

As described in section 2.3 in [I-D.ietf-bier-te-arch], BIER-TE inherits the encapsulation supporting from BIER unchanged. The encapsulation defined in [RFC8296], which specifies a common header format for both MPLS and non-MPLS networks, though the first 20-bits (referred to as BIFT-id) of the header is an "MPLS Label" in case of MPLS networks and is a local 20-bit opaque value in case of non-MPLS networks.

As described in section 4.3 of [I-D.ietf-bier-te-arch], it is necessary to distinguish the BIER and BIER-TE packet and forwarding. like [RFC8401] and [I-D.ietf-bier-lsr-non-mpls-extensions], the MPLS and non-MPLS encapsulation needs to be advertised for BIER-TE packet encapsulation.

This document describes the IS-IS protocol extension that is required for BIER-TE with MPLS and non-MPLS encapsulation associated with link BitPosition distributing.
1.1. Requirements Language

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].

2. Terminology

This document does not introduce more terminologies than [RFC8279], [RFC8296], [RFC8401], [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-isis].

3. Flooding of the BIER-TE Encapsulation Information in IS-IS

[I-D.ietf-bier-te-isis] defines a BIER-TE Info Sub-TLV to distribute the Link BitPosition information about the link. This document defines the MPLS Encapsulation Sub-sub-TLV and the Non-MPLS Encapsulation Sub-sub-TLV associate with the BIER-TE Sub-TLV to carry the MPLS and non-MPLS encapsulation information.

In case both of the two encapsulation sub-sub TLVs are advertised by one node, the label in MPLS encapsulation sub-sub TLV and BIFT-id in non-MPLS encapsulation sub-sub TLV SHOULD NOT be overlapped.

3.1. MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with [RFC8401] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.

```
| 0 | 1 | 2 | 3 |
+---+---+---+---+
| Type | Length |
| Max SI | Label |
| BS Len | Reserved |
```

Figure 1

Type: TBD1 (To be assigned by IANA).

Length: 2-octet.

Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.
Label: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The "label range" is the set of labels beginning with the Label and ending with (Label + (Max SI)). A unique label range is allocated for each BitString length and sub-domain-id.

The size of the label range is determined by the number of SIs (Section 1 of [RFC8279]) that are used in the network. Each SI maps to a single label in the label range: the first label is for SI=0, the second label is for SI=1, etc.

If the label associated with the Maximum Set Identifier exceeds the 20-bit range, the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

If the BitString length is set to a value that does not match any of the allowed values specified in [RFC8296], the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

The "BIFT-id range" is the set of 20-bit values beginning with the BIFT-id and ending with (BIFT-id + (Max SI)).

The flooding scope of this information is the same with section 5 in [RFC8401].

3.2. Non-MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with section 3.1 in [I-D.ietf-bier-lsr-non-mpls-extensions] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.
Type: TBD2 (To be assigned by IANA).

Length: 2-octet.

Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.

BIFT-id: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The BIFT-id range, limitation and flooding scope is the same with the MPLS encapsulation sub-sub-TLV.

3.2.1. BIER-TE IPv6 Encapsulation Advertisement

As described in section 4 of [I-D.ietf-bier-bierin6], which describes how the existing BIER encapsulation specified in [RFC8296] works in a non-MPLS IPv6 network, a node that requires IPv6 encapsulation MUST advertise the BIER-TE IPv6 encapsulation sub-sub-sub-TLV, which follows the Non-MPLS Encapsulation Sub-sub-TLV, according to local configuration or policy in the BIER domain to request other BFRs to always use IPv6 encapsulation.
Type: TBD3 (To be assigned by IANA).

Length: 0.

4. IANA Considerations

TBD.

5. Security Considerations

This document does not introduce more security considerations than [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-isis].

6. References

6.1. Normative References

[I-D.ietf-bier-bierin6]

[I-D.ietf-bier-lsr-non-mpls-extensions]

[I-D.ietf-bier-te-arch]
[I-D.ietf-bier-te-isis]
Chen, H., McBride, M., Wang, A., Mishra, G. S., Fan, Y.,
Liu, L., and X. Liu, "IS-IS Extensions for BIER-TE", Work
in Progress, Internet-Draft, draft-ietf-bier-te-isis-00,
25 January 2022, <https://www.ietf.org/archive/id/draft-
ietf-bier-te-isis-00.txt>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,

Przygienda, T., and S. Aldrin, "Multicast Using Bit Index
Explicit Replication (BIER)", RFC 8279,
DOI 10.17487/RFC8279, November 2017,

Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation
for Bit Index Explicit Replication (BIER) in MPLS and Non-
MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January

Zhang, "Bit Index Explicit Replication (BIER) Support via
IS-IS", DOI 10.17487/RFC8401, RFC 8401, June 2018,

Authors’ Addresses
Zheng Zhang
ZTE Corporation
China
Email: zhang.zheng@zte.com.cn

Yuehua Wei
ZTE Corporation
China
Email: wei.yuehua@zte.com.cn

Benchong Xu
ZTE Corporation
China
Email: xu.benchong@zte.com.cn

Abstract

This document describes the OSPF protocol extension that is required for BIER-TE with MPLS and non-MPLS 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 6 January 2023.

Copyright Notice

Copyright (c) 2022 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.
1.  Introduction

[I-D.ietf-bier-te-arch] describes per-packet stateless strict and loose path steered replication and forwarding for "Bit Index Explicit Replication" (BIER, [RFC8279]) packets. It is called BIER Tree Engineering (BIER-TE) and is intended to be used as the path steering mechanism for Traffic Engineering with BIER.

BIER-TE introduces a new semantic for "bit positions" (BP) of link. [I-D.ietf-bier-ospf-te] describes OSPF extensions for distributing BitPositions configured on the links in "Bit Index Explicit Replication Traffic Engineering" (BIER-TE) domain.

As described in section 2.3 in [I-D.ietf-bier-te-arch], BIER-TE inherits the encapsulation supporting from BIER unchanged. The encapsulation defined in [RFC8296], which specifies a common header format for both MPLS and non-MPLS networks, though the first 20-bits (referred to as BIFT-id) of the header is an "MPLS Label" in case of MPLS networks and is a local 20-bit opaque value in case of non-MPLS networks.

As described in section 4.3 of [I-D.ietf-bier-te-arch], it is necessary to distinguish the BIER and BIER-TE packet and forwarding. Like [RFC8444] and [I-D.ietf-bier-lsr-non-mpls-extensions], the MPLS and non-MPLS encapsulation needs to be advertised for BIER-TE packet encapsulation.

This document describes the OSPF protocol extension that is required for BIER-TE with MPLS and non-MPLS encapsulation associated with link BitPosition distributing.
1.1. Requirements Language

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].

2. Terminology

This document does not introduce more terminologies than [RFC8279], [RFC8296], [RFC8444], [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-ospf].

3. Flooding of the BIER-TE Encapsulation Information in OSPF

[I-D.ietf-bier-te-ospf] defines a BIER-TE Sub-TLV to distribute the Link BitPosition information about the link. This document defines the MPLS Encapsulation Sub-sub-TLV and the Non-MPLS Encapsulation Sub-sub-TLV associate with the BIER-TE Sub-TLV to carry the MPLS and non-MPLS encapsulation information.

In case both of the two encapsulation sub-sub-TLVs are advertised by one node, the label in MPLS encapsulation sub-sub-TLV and BIFT-id in non-MPLS encapsulation sub-sub-TLV SHOULD NOT be overlapped.

3.1. MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with [RFC8444] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.

```
0                   1                   2                   3
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Max SI    |                     Label                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|BS Len |                     Reserved                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1

Type: TBD1 (To be assigned by IANA).

Length: 2-octet.

Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.
Label: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The "label range" is the set of labels beginning with the Label and ending with (Label + (Max SI)). A unique label range is allocated for each BitString length and sub-domain-id.

The size of the label range is determined by the number of SIs (Section 1 of [RFC8279]) that are used in the network. Each SI maps to a single label in the label range: the first label is for SI=0, the second label is for SI=1, etc.

If the label associated with the Maximum Set Identifier exceeds the 20-bit range, the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

If the BitString length is set to a value that does not match any of the allowed values specified in [RFC8296], the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

The "BIFT-id range" is the set of 20-bit values beginning with the BIFT-id and ending with (BIFT-id + (Max SI)).

The flooding scope of this information is the same with section 2.3 in [RFC8444].

3.2. Non-MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with section 3.2 in [I-D.ietf-bier-lsr-non-mpls-extensions] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.
Type: TBD2 (To be assigned by IANA).

Length: 2-octet.

Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.

BIFT-id: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The BIFT-id range, limitation and flooding scope is the same with the MPLS encapsulation sub-sub-TLV.

4. IANA Considerations

TBD.

5. Security Considerations

This document does not introduce more security considerations than [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-ospf].

6. References

6.1. Normative References
Internet-Draft            Abbreviated Title            July 2022

[I-D.ietf-bier-lsr-non-mpls-extensions]
  Dhanaraj, S., Yan, G., Wijnands, I., Psenak, P., Zhang, Z., and J. Xie,
  "LSR Extensions for BIER non-MPLS Encapsulation", Work in Progress,
  Internet-Draft, draft-ietf-bier-lsr-non-mpls-extensions-00, 1 March 2022,

[I-D.ietf-bier-te-arch]
  Eckert, T., Menth, M., and G. Cauchie, "Tree Engineering for Bit
  Index Explicit Replication (BIER-TE)", Work in Progress,
  Internet-Draft, draft-ietf-bier-te-arch-13, 25 April 2022,

[I-D.ietf-bier-te-ospf]
  Chen, H., McBride, M., Wang, A., Mishra, G. S., Fan, Y., Liu, L.,
  and X. Liu, "OSPF Extensions for BIER-TE", Work in Progress,
  Internet-Draft, draft-ietf-bier-te-ospf-00, 20 January 2022,

[RFC2119]  Bradner, S., "Key words for use in RFCs to Indicate
  Requirement Levels", BCP 14, RFC 2119,
  DOI 10.17487/RFC2119, March 1997,

  Przygienda, T., and S. Aldrin, "Multicast Using Bit Index
  Explicit Replication (BIER)", RFC 8279,
  DOI 10.17487/RFC8279, November 2017,

  Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for
  Bit Index Explicit Replication (BIER) in MPLS and Non-
  MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January

[RFC8444]  Psenak, P., Ed., Kumar, N., Wijnands, IJ., Dolganow, A.,
  Przygienda, T., Zhang, J., and S. Aldrin, "OSPFv2
  Extensions for Bit Index Explicit Replication (BIER)",
  RFC 8444, DOI 10.17487/RFC8444, November 2018,

Authors’ Addresses

OSPFv3 extensions for BIER-TE (Tree Engineering for Bit Index Explicit Replication) with MPLS and non-MPLS Encapsulation
draft-zwx-bier-te-ospfv3-extensions-01

Abstract

This document describes the OSPFv3 protocol extension that is required for BIER-TE with MPLS and non-MPLS 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 25 January 2023.

Copyright Notice

Copyright (c) 2022 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.
1. Introduction

[I-D.ietf-bier-te-arch] describes per-packet stateless strict and loose path steered replication and forwarding for "Bit Index Explicit Replication" (BIER, [RFC8279]) packets. It is called BIER Tree Engineering (BIER-TE) and is intended to be used as the path steering mechanism for Traffic Engineering with BIER.

BIER-TE introduces a new semantic for "bit positions" (BP) of link. [I-D.ietf-bier-te-ospfv3] describes OSPFv3 extensions for distributing BitPositions configured on the links in "Bit Index Explicit Replication Traffic Engineering" (BIER-TE) domain.

As described in section 2.3 in [I-D.ietf-bier-te-arch], BIER-TE inherits the encapsulation supporting from BIER unchanged. The encapsulation defined in [RFC8296], which specifies a common header format for both MPLS and non-MPLS networks, though the first 20-bits (referred to as BIFT-id) of the header is an "MPLS Label" in case of MPLS networks and is a local 20-bit opaque value in case of non-MPLS networks.

As described in section 4.3 of [I-D.ietf-bier-te-arch], it is necessary to distinguish the BIER and BIER-TE packet and forwarding. like [I-D.ietf-bier-ospf-extensions] and [I-D.ietf-bier-lsr-non-mpls-extensions], the MPLS and non-MPLS encapsulation needs to be advertised for BIER-TE packet encapsulation.

This document describes the OSPFv3 protocol extension that is required for BIER-TE with MPLS and non-MPLS encapsulation associated with link BitPosition distributing.
1.1. Requirements Language

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].

2. Terminology

This document does not introduce more terminologies than [RFC8279], [RFC8296], [I-D.ietf-bier-ospfv3-extensions], [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-ospfv3].

3. Flooding of the BIER-TE Encapsulation Information in OSPFv3

[I-D.ietf-bier-te-ospfv3] defines a BIER-TE Sub-TLV to distribute the Link BitPosition information about the link. This document defines the MPLS Encapsulation Sub-sub-TLV and the Non-MPLS Encapsulation Sub-sub-TLV associate with the BIER-TE Sub-TLV to carry the MPLS and non-MPLS encapsulation information.

In case both of the two encapsulation sub-sub-TLVs are advertised by one node, the label in MPLS encapsulation sub-sub-TLV and BIFT-id in non-MPLS encapsulation sub-sub-TLV SHOULD NOT be overlapped.

3.1. MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with [I-D.ietf-bier-ospfv3-extensions] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.

```
+----------------+----------------+----------------+----------------+
|       Type     |      Length    |
+----------------+----------------+
|     Max SI     |       Label    |
+----------------+----------------+
|      BS Len    |     Reserved   |
+----------------+----------------+
```

Figure 1

Type: TBD1 (To be assigned by IANA).

Length: 2-octet.
Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.

Label: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The "label range" is the set of labels beginning with the Label and ending with (Label + (Max SI)). A unique label range is allocated for each BitString length and sub-domain-id.

The size of the label range is determined by the number of SIs (Section 1 of [RFC8279]) that are used in the network. Each SI maps to a single label in the label range: the first label is for SI=0, the second label is for SI=1, etc.

If the label associated with the Maximum Set Identifier exceeds the 20-bit range, the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

If the BitString length is set to a value that does not match any of the allowed values specified in [RFC8296], the BIER MPLS Encapsulation Sub-TLV containing the error MUST be ignored.

The "BIFT-id range" is the set of 20-bit values beginning with the BIFT-id and ending with (BIFT-id + (Max SI)).

The flooding scope of this information is the same with section 2.2 in [I-D.ietf-bier-ospfv3-extensions].

3.2. Non-MPLS Encapsulation Sub-sub-TLV

The following description is the almost same with section 3.3 in [I-D.ietf-bier-lsr-non-mpls-extensions] since [I-D.ietf-bier-te-arch] inherit the encapsulation of BIER header.
Type: TBD2 (To be assigned by IANA).

Length: 2-octet.

Max SI: A 1-octet field encoding the maximum Set Identifier (SI) (see Section 1 of [RFC8279]) used in the encapsulation for this BIER sub-domain for this BitString length.

BIFT-id: A 3-octet field, where the 20 rightmost bits represent the first label in the label range. The 4 leftmost bits MUST be ignored.

BS Len (BitString Length): A 4-bit field encoding the supported BitString length associated with this BFR-prefix. The values allowed in this field are specified in Section 2 of [RFC8296].

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The BIFT-id range, limitation and flooding scope is the same with the MPLS encapsulation sub-sub-TLV.

3.2.1. BIER-TE IPv6 Encapsulation Advertisement

As described in section 4 of [I-D.ietf-bier-bierin6], which describes how the existing BIER encapsulation specified in [RFC8296] works in a non-MPLS IPv6 network, a node that requires IPv6 encapsulation MUST advertise the BIER-TE IPv6 encapsulation sub-sub-sub-TLV, which follows the Non-MPLS Encapsulation Sub-sub-TLV, according to local configuration or policy in the BIER domain to request other BFRs to always use IPv6 encapsulation.
4. IANA Considerations

TBD.

5. Security Considerations

This document does not introduce more security considerations than [I-D.ietf-bier-te-arch] and [I-D.ietf-bier-te-ospfv3].

6. References

6.1. Normative References

[I-D.ietf-bier-bierin6]

[I-D.ietf-bier-lsr-non-mpls-extensions]

[I-D.ietf-bier-ospfv3-extensions]
[I-D.ietf-bier-te-arch]

[I-D.ietf-bier-te-ospfv3]


Authors’ Addresses

Zheng Zhang
ZTE Corporation
China
Email: zhang.zheng@zte.com.cn

Yuehua Wei
ZTE Corporation
China
Email: wei.yuehua@zte.com.cn

Benchong Xu
ZTE Corporation
China

Abstract

Bit Index Explicit Replication (BIER) is a multicast technology with a new encapsulation and forwarding paradigm. BIER encapsulation is specified in RFC8296, and this document specifies extension headers used with BIER encapsulation header.

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 12 January 2023.

Copyright Notice

Copyright (c) 2022 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
1. Background

[RFC8296] specifies BIER encapsulation header as following:

```
+-------------------+-------------------+-------------------+-------------------+-------------------+
| BIFT-id           | TC                | S                | TTL               |
+-------------------+-------------------+-------------------+-------------------+
| Nibble | Ver | BSL | Entropy              |
+-------------------+-------------------+-------------------+-------------------+
| OAM | Rsv | DSCP | Proto | BFIR-id            |
+-------------------+-------------------+-------------------+-------------------+
| BitString (first 32 bits) |
+-------------------+-------------------+-------------------+-------------------+
| BitString (last 32 bits) |
+-------------------+-------------------+-------------------+-------------------+
```

The header is fixed format, making it difficult to extend for additional functionalities.

[I-D.xzlnp-bier-ioam] describes In-situ Operations, Administration, and Maintenance (IOAM) functionality for BIER, and proposes a way to extend the BIER header to encode IOAM data as well as other potential extension headers.
[I-D.zzhang-intarea-generic-delivery-functions] considers Generic Delivery Functions (GDFs, e.g., fragmentation, security, IOAM) that can be applied to various layers (e.g., MPLS, IPv6, BIER) and proposes a slightly different extension header mechanism that is aligned with [I-D.song-mpls-extension-header] and work for both MPLS, BIER, and potentially other layers. This approach is adopted in this BIER document for formal discussions and progressing in the BIER WG.

2. Specifications

2.1. BIER Encapsulation with Extension Headers

The following figure illustrates a BIER header with extension headers.

A TBD value for the "proto" field in the BIER header indicates that some BIER Extension Headers follow the BIER header and precede the BIER payload. An extension header could be an IPv6 one for a GDF like IOAM, even though this is not an IPv6 protocol layer.

R: The "R" flag bit is reserved. It MUST be set to 0 on transmission and ignored on reception.
If the "H" flag bit is set, it indicates the presence of at least one extension header that needs to be processed hop by hop even before a BFER is reached.

2.2. Head of Extension Headers

The Header of Extension Headers (HEH) is specified in [I-D.song-mpls-extension-header] and quoted here:

```
 0123 4567 89012345 67890123 45678901
+----+----+--------+--------+--------+
 | R  |EHC |  EHTL  |  OUL   |   NH   |
+----+----+--------+--------+--------+
```

The meaning of the fields in an HEH is as follows:

**R**: 4-bit reserved. The nibble value means to avoid conflicting with IP version numbers.

**EHC**: 4-bit unsigned integer for the Extension Header Counter. This field keeps the total number of extension headers included in this packet. It does not count the original upper layer protocol headers. At most 15 EHs are allowed in one packet.

**EHTL**: 8-bit unsigned integer for the Extension Header Total Length in 4-octet units. This field keeps the total length of the extension headers in this packet, not including the HEH itself.

**OUL**: 8-bit Original Upper Layer protocol number indicating the original upper layer protocol type. It can be set to "UNKNOWN" if unknown.

**NH**: 8-bit selector for the Next Header. This field identifies the type of the header immediately following the HEH. Value is from the Internet Protocol Numbers registry, so that IPv6 extension headers for GDFs can be used as BIER extension headers as is. Only one new TBD value needs to be assigned from the registry for extension headers specific for all non-IPv6 layers that use this common extension header design. All those layers, e.g., BIER and MPLS, upon encounter an extension header referred to by the TBD NH will process it as specific to that layer and according to its EXT field (see section below).
The 4-bit R field of HEH is not important to BIER, but it’s used for other layers like MPLS. The NH field in HEH and EH aligns this with IPv6 extension headers - applicable IPv6 extension headers can be used here as is (the design goal is to allow generic functions to be applied at different layers with the same extension header format).

2.3. Extension Header Format

The format of a layer specific Extension Header (EH) is as following:

```
0          1          2          3
01234567 89012345 67890123 45678901
+--------+--------+--------+-------+
<table>
<thead>
<tr>
<th>NH</th>
<th>HLEN</th>
<th>EXT</th>
<th>resved</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>˜        Header Specific Data      ˜</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------+--------+----------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+--------+--------+----------------+
```

It is like an IPv6 extension header with an added EXT field.

NH: 8-bit indicator for the Next Header. This field identifies the type of the EH immediately following this EH.

HLEN: 8-bit unsigned integer for the Extension Header Length in 8-octet units, not including the first 8 octets.

EXT: 8-bit extension type (sub-type) for BIER specific extension headers referred to by the TBD NH mentioned above. Value is allocated from a registry specifically for BIER extension header. It is expected to have different EXT values for different BIER specific extension headers for different functions, and an EXT value for one extension header that encodes different TLVs for different functions in a single EH.

Header Specific Data: Variable length field for the specification of the EH.

3. Security Considerations

To be provided.

4. IANA Considerations

The following IANA actions are requested.
Assign a "BIER Extension Header" value from the "BIER Next Protocol Identifiers" registry for the "proto" field in a BIER header.

Assign a TBD value from "IP Protocol Numbers" registry for "Non-IP Extension Header".

Create a "BIER Extension Header Sub-type" registry for the "EXT" field in a BIER specific extension header.

Assign a "TLV-encoded BIER Extension Header" value from the "BIER Extension Header Sub-type" registry.

5. Acknowledgements

6. References

6.1. Normative References

[I-D.song-mpls-extension-header]


6.2. Informative References

[I-D.xzlnp-bier-ioam]
Zhang, Z., Bonica, R., Kompella, K., and G. Mirsky,
"Generic Delivery Functions", Work in Progress, Internet-Draft, draft-zzhang-intarea-generic-delivery-functions-03,

Authors’ Addresses

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net

Xiao Min
ZTE
Email: xiao.min2@zte.com.cn

Yisong Liu
China Mobile
Email: liuyisong@chinamobile.com

Hootman Bidgoli
Nokia
Email: hooman.bidgoli@nokia.com