BIER Prefix Redistribute

Abstract

This document defines a BIER proxy function to support a single BIER
sub-domain over multiple underlay routing protocol regions
(Autonomous Systems or IGP areas). A new BIER proxy range sub-TLV is
defined to redistribute BIER BFR-id information across the routing
regions.

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

Copyright Notice

Table of Contents

1. Introduction

[RFC8279] introduces a novel multicast architecture. It does not
require a signaling protocol to explicitly build multicast
distribution trees, nor does it require intermediate nodes to
maintain any per-flow state.

OSPF/ISIS/BGP signaling for BIER [RFC8401], [RFC8444],
[I-D.ietf-bier-idr-extensions], [I-D.ietf-bier-ospfv3-extensions],
[I-D.ietf-bier-lsr-non-mpls-extensions] define the extensions to
support BIER information distribution in BIER domains.

This document defines a BIER proxy function to support a single BIER
sub-domain over multiple underlay routing protocol regions
(Autonomous Systems or IGP areas).

2. Terminology

The terminologies are the same with the definition in BIER architecture [RFC8279], OSPF/ISIS/BGP signaling for BIER [RFC8401], [RFC8444], [I-D.ietf-bier-idr-extensions], [I-D.ietf-bier-ospfv3-extensions].

3. Problem statement

BIER [RFC8279] is a new multicast architecture that does not need per-tree state inside a network for multicast forwarding. BIER forwarding state (which is not per-tree) is built according to a routing underlay, which is defaulted to an IGP domain though not limited to that. In this routing underlay, BIER information like BIER-prefix, BFR-id, subdomain-id, and encapsulation is propagated using IGP or BGP as specified in documents such as [RFC8401], [RFC8444], [I-D.ietf-bier-idr-extensions], [I-D.ietf-bier-ospfv3-extensions], [I-D.ietf-bier-lsr-non-mpls-extensions].

In some deployment situations, different routing protocols may be used in differet parts of a network yet there are just small number of BFERs in each protocol domain, so a single BIER sub-domain is desired. This requires BIER information redistribution among different regions or protocols, as described in the following two sections.

3.1. Multipe IGP domains

```
              +----+               +----+
        +----+ R1 +-------------+ R2 +-----+
        |    +----+             +----+     |
        |            OSPF / ISIS           |
        |             domain 1             |
        |                                  |
        |     +----+             +----+    |
        +-----+    +-----------+     +-----+
    +-----------+ R3 +---+    +---+ R4 +-----------+
    |            +----+   |    |   +----+          |
    |        OSPF         |    |        OSPF       |
    |                     |    |                   |
    |     domain 2        |    |     domain 3      |
    | +----+     +----+   |    |   +----+    +----+ |
    +--+ Rm +-----+ Rn +--+    +---+ Rx +-----+ Ry +--+
       +----+     +----+           +----+    +----+
```
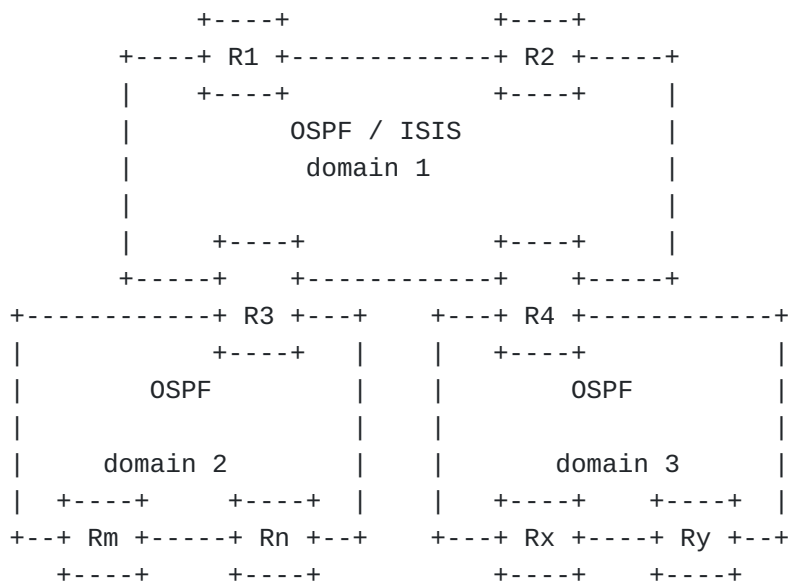   Figure 1

While one could treat each IGP domain in the above network as a separate BIER sub-domain, the border routers R3/R4 would need

decapsulate incoming BIER header in one BIER sub-domain, forward
based on flow overlay per-tree state, and re-ecapsulate with a BIER
header for forwarding in the next BIER sub-domain. This not only is
inefficient in forwarding, but also require per-tree state on the
border routers, which is undesired.

A better solution is to treat the entire network of multiple IGP
domains as single BIER sub-domain with a single routing underlay.

3.2. Seamless MPLS Network

Figure 1 could also depict a Seamless MPLS
[I-D.ietf-mpls-seamless-mpls] network, where BGP-LU [RFC8277] is
used to distribute routes for edge routers (e.g. R1/R2/Rm/Rn/Rx/Ry)
among the edge routers and border routers (e.g. R3/R4), but those
routes are not redistributed into other IGP areas/domains. With
that, internal routers in an area/domain will only have routes to
local nodes, yet they still need to build BIER forwarding state for
BFERs in other areas/domains.

4. Proposed Solution

For the multiple IGP domains scenario in Section 3.1, BIER
information from one domain needs to be redistributed into another
domain, like that BIER information is redistributed from one IGP
area to another as specified in [RFC8401], [RFC8444], and
[I-D.ietf-bier-ospfv3-extensions].

Specifically, when an ASBR redistributes BIER prefixes for BFERs
from one protocol domain to another, BIER information is also
redistributed except the encapsulation information (because BFRs in
one domain will not directly send BIER packets to BFRs in the other
domain so only the BFR-IDs of the BFERs matters). When BIER prefixes
for non-BFIR/BFER (i.e. whose BFR-ID is 0) are redistributed, BIER
information is not redistributed.

If route summarization is used, because a summarized prefix may
cover many BFERs, the BFR-IDs of those covered BFERs needs to be
explicitly listed in proxy range sub-TLV (see Section 5.2). In case
of Seamless MPLS (Section 3.2), when a border router advertise BIER
information for itself in one area/domain, it also explicitly lists
the BFIRs/BFERs in other areas/domains that are reachable via itself
in the proxy range sub-TLV.

In figure 1, R3/R4 connects two routing domains. After R3 receives
BIER information for Rm/Rn from domain 2 and redistribute to domain
1, BFRs in domain 1 can build BIER forwarding state for BFERs in
domain 2 through R3. Similarly, R3 receives BIER information for R1/
R2 from domain 1 and redistribute into domain 2. BFRs in domain 2
can build BIER forwarding state for BFERs in domain 1 through R3.

For example, in this network, suppose that Rm and Rn have the prefix of 203.0.113.1/32, 203.0.113.2/32. In order to build one BIER sub-domain which includes these three IGP domains, R3 advertises the BFR-ids of Rm/Rn with associated prefixes (203.0.113.1/32, 203.0.113.2/32) into the upper domain. Similarly, R4 advertises the BFR-ids of Rx/Ry with associated prefixes (198.51.100.1/32, 198.51.100.2/32) into the upper domain too.

And R3/R4 advertises the prefixes of R1 and R2 (suppose that the prefixes are 192.0.2.1/32 and 192.0.2.2/32) with associated BFR-ids into IGP domain 1 and domain 2. Also, R3 advertises the prefixes learned from R4 (198.51.100.1/32, 198.51.100.2/32) with associated BFR-ids into IGP domain 1. R4 also advertises the prefixes (203.0.113.1/32, 203.0.113.2/32) with associated BFR-ids into IGP domain 2.

Obviously, in order to build the large BIER sub-domain, the BFR-id of edge router in each IGP domain MUST NOT overlap.

5. Specifications

5.1. Redistribution of BIER Information

Consider a BIER sub-domain that spans multiple routing domains. The procedures in this section apply if a border router, which is also a BFR, redistribute routing information from one routing domain into another.

If a redistributed route is for a host route for a BFIR/BFER (i.e. the BFR-ID is not zero) in the same sub-domain, BIER information for the BFIR/BFERs MUST be advertised in the target routing domain as following:

  *If the target routing domain is OSPFv2, a BIER Sub-TLV [RFC8444] is attached to the OSPFv2 Extended Prefix TLV in the OSPFv2 Extended Prefix Opaque LSA [RFC7684] corresponding to the redistributed host route.

  *If the target routing domain is OSPFv3, a BIER Sub-TLV [I-D.ietf-bier-ospfv3-extensions] is attached to the OSPFv3 Extended LSA TLVs in the Intra-Area-Prefix TLV [RFC8362]. corresponding to the redistributed host route.

  *If the target routing domain is ISIS, a BIER Info Sub-TLV [RFC8401] is attached to the TLVs of 235, 237, [RFC5120], 135 [RFC5305], or 236 [RFC5308].

  *If the target routing domain is BGP, a BIER TLV [I-D.ietf-bier-idr-extensions] is attached to BGP Path Attribute.

The BIER Sub-TLV (in case of OSPF2 and OSPFv3), BIER Info Sub-TLV
(in case of IS-IS) and BIER TLV (in case of BGP) are encoded as
specified in [RFC8444], [I-D.ietf-bier-ospfv3-extensions], and
[RFC8401], and [I-D.ietf-bier-idr-extensions]. The encapsulation
sub-TLV SHOULD NOT be included because it would not be used.

If a redistributed route is for a host route for a transit BFR (i.e.
the BFR-ID is zero), BIER information for the BFR SHOULD NOT be
redistributed, because it would not be used.

If the redistributed route is a summary or default route that covers
some BFIR/BFERs, a BIER Sub-TLV (for OSPFv2 and OSPFv3), or a BIER
Info Sub-TLV (in case of IS-IS), or a BIER TLV (in case of BGP) MUST
be used to advertise the covered BFIRs/BFERs via the BIER proxy
range sub-TLV as specified in the following section. The proxy range
sub-TLV MAY also be used when the BIER prefix is for a border router
via which multiple BFERs can be reached.

5.2. BIER proxy range sub-TLV

The BIER Sub-TLV can include a proxy range sub-TLV, which lists
BFIRs/BFERs covered by a prefix or a summary/default route or
reachable via a BFR.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |             resv              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |          BFR-id             |            BFR-id range          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                            ...                                |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |          BFR-id             |            BFR-id range          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                            figure 2
```

  *Type: 8-bit unsigned integer. TBD to indicate the BIER proxy
   range sub-TLV.

  *Length: 8-bit unsigned integer. Length of the BIER proxy range
   sub-TLV in 4-octet units, not including the first 4 octets.

  *resv: 16-bit unsigned integer. The reserved field.

  *BFR-id: 16-bit unsigned integer. The first BFR-id from original
   advertisement.

  *BFR-id range: 16-bit unsigned integer. The range of BFR-ids with
   one subdomain-id.

The BIER proxy range sub-TLV is included in the BIER Sub-TLV for an aggregated/summary route prefix or default route prefix, or in the BIER Sub-TLV for a BIER prefix (i.e., a border router in a Seamless MPLS network). Multiple BIER proxy range sub-TLVs MAY be included in the BIER Sub-TLV.

The range in the BIER proxy range sub-TLV can be as granular as to advertise individual BFR-ids. Though a larger range can increase advertisement efficiency, that requires careful planning for BFR-id assignment.

When the proxy range sub-TLV is used, the mapping between a BIER prefix and its BFR-id is no longer conveyed in the routing underlay. As a result, the mapping must be provided by other means, e.g. in the multicast overlay.

5.3. BIRT/BIFT Calculation

If a BFR receives a BIER prefix whose BIER Sub-TLV includes a proxy range sub-TLV (i.e., the Seamless MPLS scenario), it treats as if that the originator advertised a default route with the proxy range sub-TLV. Note that this imaginary default route is only for the purpose of building BIRT/BIFT entries and not used for unicast forwarding.

With the BIER prefixes originated in the local routing area/domain, the BIER prefixes and summary/default routes redistributed into the local routing area/domain, and the imaginary default route mentioned above, a BFR builds BIRTs as specified in [RFC8279] with entries including host/summary/default prefixes.

BIFT entries are then derived from a corresponding BIRT. For a BFER covered by the proxy range sub-TLVs associated with the summary/ default prefixes (whether or not the deafult prefix is the imaginary one as mentioned above), its BIFT entry is derived from the summary/ default prefix entry in the BIRT. It is possible that a BFR-ID for a BFER is listed in the proxy range sub-TLV of multiple prefixes. if one prefix is less specific than another, it is not considered for the BFER. Of all the remaining prefixes whose proxy range sub-TLV covers a BFR-ID, the one with the preferred cost/metric MUST be used to derive the BIFT entry for the BFER. When there is a tie, ECMP or tie-breaker MAY be used.

With this scheme, even though the BIER prefixes are not advertised into the IGP for an area/domain in a Seamless MPLS network and unicast traffic for those BIER prefixes are tunneled through, corresponding BIFT entries are maintained inside the area/domain for the purpose of efficient BIER forwarding. Otherwise, BIER forwarding through the area/domain would be tunneled just like unicast case.

6. IANA Considerations

IANA is requested to set up a new types of sub-TLV (TLV) registry value for BIER proxy range advertisement in OSPF, ISIS, BGP, etc.

7. Security Considerations

Implementations must assure that malformed TLV and Sub-TLV permutations do not result in errors which cause hard protocol failures.

8. Acknowledgements

The authors would like to thank Stig Venaas for his valuable comments and suggestions.

9. References

9.1. Normative References

[I-D.ietf-bier-idr-extensions]
Xu, X., Chen, M., Patel, K., Wijnands, I., Przygienda, T., and Z. J. Zhang, "BGP Extensions for BIER", Work in Progress, Internet-Draft, draft-ietf-bier-idr-extensions-09, 15 February 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bier-idr-extensions-09>.

[I-D.ietf-bier-lsr-non-mpls-extensions]
Dhanaraj, S., Yan, G., Wijnands, I., Psenak, P., Zhang, Z. J., and J. Xie, "LSR Extensions for BIER non-MPLS Encapsulation", Work in Progress, Internet-Draft, draft-ietf-bier-lsr-non-mpls-extensions-01, 19 September 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-bier-lsr-non-mpls-extensions-01>.

[I-D.ietf-bier-ospfv3-extensions] Psenak, P., Nainar, N. K., and I. Wijnands, "OSPFv3 Extensions for BIER", Work in Progress, Internet-Draft, draft-ietf-bier-ospfv3-extensions-07, 1 December 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-bier-ospfv3-extensions-07>.

[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>.

[RFC8279]   Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A.,
            Przygienda, T., and S. Aldrin, "Multicast Using Bit Index
            Explicit Replication (BIER)", RFC 8279, DOI 10.17487/
            RFC8279, November 2017, <https://www.rfc-editor.org/info/rfc8279>.

[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>.

[RFC8401]   Ginsberg, L., Ed., Przygienda, T., Aldrin, S., and Z.
            Zhang, "Bit Index Explicit Replication (BIER) Support via
            IS-IS", RFC 8401, DOI 10.17487/RFC8401, June 2018,
            <https://www.rfc-editor.org/info/rfc8401>.

[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, <https://www.rfc-editor.org/info/rfc8444>.

9.2. Informative References

[I-D.ietf-mpls-seamless-mpls]
            Leymann, N., Decraene, B., Filsfils, C., Konstantynowicz,
            M., and D. Steinberg, "Seamless MPLS Architecture", Work
            in Progress, Internet-Draft, draft-ietf-mpls-seamless-mpls-07, 28 June 2014, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-seamless-mpls-07>.

[RFC8277]   Rosen, E., "Using BGP to Bind MPLS Labels to Address
            Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017,
            <https://www.rfc-editor.org/info/rfc8277>.

Appendix A. Proxy range sub-TLV usage

   This appendix is to make the function understood more easily. Except
   for inter-area case, the function is also suitable for inter-as
   case. In the same example of figure 1, in case there are 40 edge
   routers in domain 1, the BFR-ids of domain 1 start from 51 to 90,
   and the prefixes of these routers start from 203.0.113.1/32 to
   203.0.113.40/32. These assigned BFR-IDs are not overlapped with the
   BFR-IDs in any other domain.

   In order to build a BIER sub-domain across these areas, the two
   advertisement methods defined in Section 5.2 can be used:

     *As a transit node, R3 advertises the BIER sub-TLV with BFR-ID set
      to 0. When R3 is not allowed to advertise the summary or specific
      prefixes into the upper domain, R3 can advertise the proxy range
      sub-TLV with the host prefix directly. So there are two sub-TLVs
      advertisement associated with the host prefix of R3.

     *As a transit node, R3 advertises the BIER sub-TLV with BFR-ID set
      to 0. When R3 is allowed to advertise the summary prefix into the
      upper domain, R3 can advertise the proxy range sub-TLV with the
      summarized prefix, 203.0.113.0/24, with the BFR-id set to 51, the
      BFR-id range set to 40, into the upper domain. In this case,
      there are two prefixes advertised by R3. But the summary prefix
      can not be used to encapsulate BIER packet directly in case of
      tunneling case. The summary prefix is only used to generate the
      BIFT.

   Then the router in the uppler domain can build the BIER forwarding
   table, the nexthops of BFR-IDs in the proxy range sub-TLV are set to
   R3.

   The same function is also applied to R4 when it advertises the BFR-
   IDs to the upper domain. This method is also applied to R3/R4 when
   it advertises the BFR-IDs to the lower domain. When R3 advertises
   the prefixes from the upper domain and domain 2 into domain 1,
   except the host prefix of R3 with BFR-ID set to 0, R3 may advertise
   only one default route (0.0.0.0/0) into domain 1 if one or more
   continuous BFR-id range can be attached. Suppose that the BFR-id in
   the upper domain starts from 1001 to 1050, the BFR-id in domain 2
   starts from 201 to 250, and these ranges are not overlapped with the
   ranges in any other domain. Suppose that the sub-domain ID is 1, the
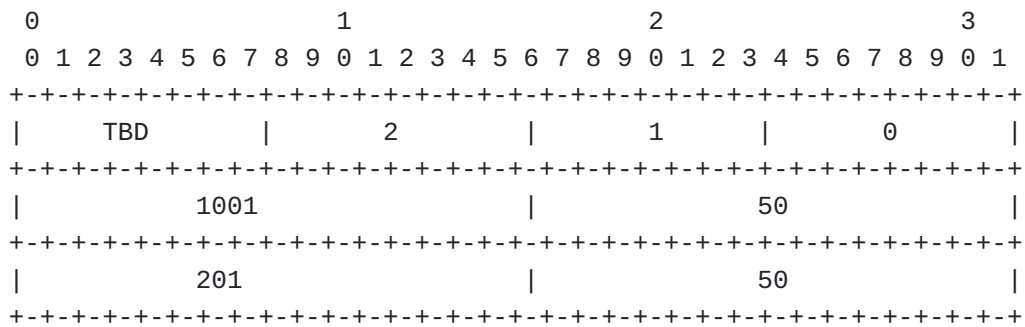   BIER proxy range sub-TLV may be advertised like this:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      TBD      |       2       |       1       |       0       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |            1001               |               50              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |            201                |               50              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                         figure 3
```

Then the BIER overlay is built among R1, R2, Rm, Rn, Rx and Ry. R3
and R4 need not maintain the multicast overlay states. The optimized
summary/ aggregated or default prefix can be generated by the
operation policy which is configured by the network administrator.

If two or more ABRs in one domain are used to reistribute the
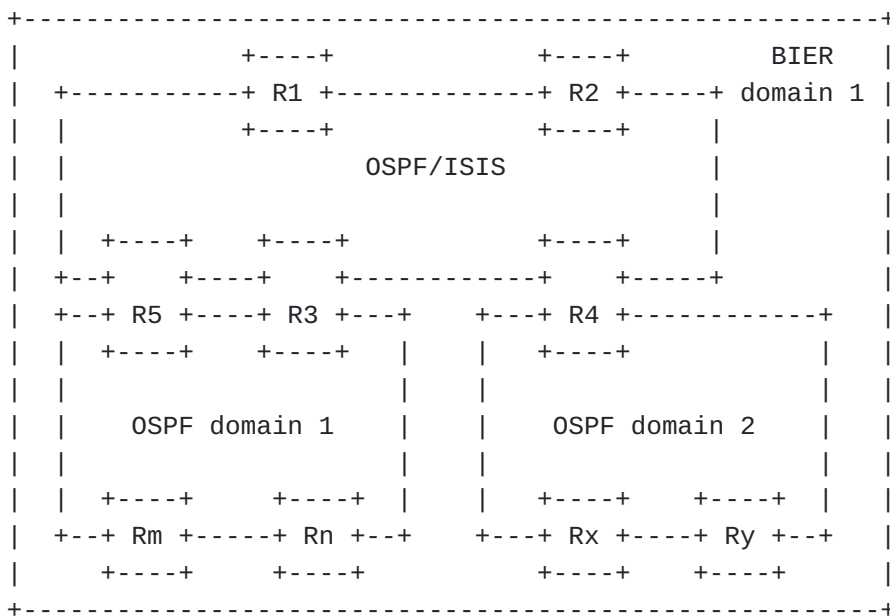prefix, for example in figure 4 below:

```
 +------------------------------------------------------------+
 |                 +----+                +----+      BIER      |
 |  +-----------+ R1 +-------------+ R2 +-----+ domain 1 |
 |  |              +----+                +----+          |     |
 |  |                       OSPF/ISIS                    |     |
 |  |                                                    |     |
 |  |    +----+     +----+                +----+         |     |
 |  +--+     +----+       +------------+     +-----+           |
 |  +--+ R5 +----+ R3 +---+     +---+ R4 +-------------+   |
 |  |  +----+      +----+    |   |    +----+                | |
 |  |  |                     |   |                        | |
 |  |      OSPF domain 1     |   |   OSPF domain 2    |   | |
 |  |  |                     |   |                    |   | |
 |  |  +----+      +----+    |   |   +----+     +----+ |   | |
 |  +--+ Rm +------+ Rn +--+     +---+ Rx +----+ Ry +--+   |
 |     +----+      +----+           +----+     +----+      |
 +------------------------------------------------------------+
                         Figure 4
```

As the ABRs, except the host prefixes of R3 and R5 advertisement, R3
and R5 can both advertise the proxy range sub-TLV with their host
prefix, the routers can select one of them as the nexthop, or select
both of them for ECMP.

If R3 and R5 are allowed to advertise the summary prefix received
from the upper domain. They can advertise the same summary prefixes
or the different prefixes according to the operation policy. When
they advertise the same summary prefixes, the R3 and R5 can also be
used for ECMP. When they advertise the different summary prefixes,
the more specific prefixes are used to generate the BIER forwarding

table. Whatever the same or different prefixes are advertised, the
nexthop is set to R3/R5.

In case the range of BFR-ids in one domain is overlapped with the
BFR-ids in any other domain, the proxy range sub-TLV may not be
used. In the same example above, if the BFR-ids in domain 1 are 21,
31, 41, etc., the BFR-ids in domain 2 are 22, 32, 42, etc., even if
the summarized prefixes are not overlapped with the prefixes in any
other domain, when R3 advertises the summarized prefixes in domain 1
into the upper domain, the proxy range sub-TLV may not optimize the
advertisement.

After the forwarding plane is built, the nexthop of the range BFR-
IDs is set to the ABR router. For example, when R1 receives
multicast packet, and the receivers of this flow are connected by Rm
and Rx, R1 encapsulates BIER header in front of the flow packet with
BFR-ids set to Rm and Rx. The routers in the upper domain forward
the packet to the ABR routers: R3, R4 and R5. When R3/R4/R5 receives
the packet, R3/R4/R5 needs not decapsulate and re-encapsulate the
packet. R3/R4/R5 just forwards the packet according to the BIER
forwarding table. Similar with the upper domain, the routers in
lower domain forward the packet to the edge routers: Rm and Rx. When
the packet reaches Rm and Rx, Rm and Rx remove the BIER header and
forward it to receivers.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation

Email: zhang.zheng@zte.com.cn

Bo Wu
Individual

Email: w1973941761@163.com

Zhaohui Zhang (editor)
Juniper Networks

Email: zzhang@juniper.net

IJsbrand Wijnands
Individual

Email: ice@braindump.be

Yisong Liu
China Mobile

        Email: liuyisong.ietf@gmail.com

   Hooman Bidgoli
   Nokia

        Email: hooman.bidgoli@nokia.com