# BIER Multicast Technical White Paper

© ZTE Corporation. All rights reserved.

# BIER Multicast Technical White Paper

| Version | **Date** | **Author** | **Reviewer** | **Remark** |
|---|---|---|---|---|
| V1.0 | 2020/08/10 | ZTE | ZTE | New |
| V1.1 | 2021/03/01 | ZTE | ZTE | New |
| | | | | |
| | | | | |

**Keywords: Stateless multicast**

Abstract: As a new stateless multicast technology, the BIER protocol keeps multicast state information on the ingress and egress nodes of multicast traffic. The intermediate nodes do not perceive multicast traffic, establish multicast forwarding trees, or maintain multicast forwarding state information. BIER multicast is applicable to large-scale multicast service deployment scenarios, such as multicast VPN service and IPTV/OTT service. BIER multicast elements carry key information such as BFR-ID, SD, BSL and encapsulation by BFR-prefix. BFR-Prefix implements network-wide flooding through traditional IGP protocols such as ISIS/OSPFv2/OSPFv3. Each BIER router in the network establishes a BIER forwarding table based on these key information to forward BIER encapsulation packets.

# Contents

# Figures

# 1. Technical Limitations of Traditional Stateful Multicast

## 1.1. Traditional Stateful Multicast Does not Adapt to Large-Scale Multicast Applications

Traditional multicast protocols such as PIM-SM/PIM-DM create a multicast distribution tree from the source to the receiver for each Group. Each node (router) in the multicast distribution tree maintains the multicast forwarding state information: (Group, Ingress interface and Egress interface). In an IPTV system, a multicast group corresponds to a TV channel, and a large-scale IPTV system supports hundreds or even thousands of subscribed channels. The traditional multicast routing protocol establishes a multicast distribution tree for each multicast group. Each router in the network maintains hundreds to thousands of multicast forwarding state information, consuming the valuable resources of routers such as CPU and TCAM. The old equipment in the existing network may face considerable pressure. When a multicast subscriber or network topology changes, the IGP protocol will reconverge. The multicast protocol can only converge again after the IGP protocol converges, and then the multicast distribution tree of each group can be calculated again. The convergence time of the multicast distributed tree is much longer than that of the IGP protocol.

## 1.2. Traditional Stateful Multicast Does not Adapt to Deployment of Large-Scale VPN Multicast

Large operators provide virtual private network services to different customers through MPLS L3VPN. Each VPN has independent address space, and the customers run independent unicast and multicast protocols such as ISIS and PIM. However, an operator cannot provide large-scale multicast services for all VPN customers because the intermediate P devices of the operator cannot maintain the multicast forwarding state of each VPN for each customer. If there are N VPN customers and each VPN customer has N multicast services, the P router of the operator needs to maintain $N^2$ multicast flow state information. If the operator's network topology is changed or the multicast source or multicast receiver is changed, each P router needs to recalculate the multicast publishing tree of each multicast group in each VPN. As a result, the convergence of VPN multicast routes is slow, and slower than that of global multicast routes, seriously affecting the multicast VPN service experience. If the operator's network topology changes or the multicast source or multicast receiver changes, each P router needs to recalculate the multicast distribution tree of each group in each VPN. As a result, the convergence of VPN multicast route is slower than that of global multicast route, which seriously affects the multicast VPN service experience. To reduce the number of VPN multicast state information maintained by P devices, operators have used various VPN multicast technologies, but have not totally solved the problem. In the multicast header replication technology used by the operators, the P node has no multicast state information, but it requires large bandwidth and high performance of the multicast PE node. The VPN multicast aggregation technology used by the operator can reduce the multicast state information of the P node, but the VPN multicast route is not optimal, which wastes the bandwidth of the WAN.

## 1.3. Traditional Stateful Multicast Does Not Conform to Network Simplification Evolution Trend

Traditional stateful multicast uses mLDP, P2MP and MP2MP technologies to forward multicast traffic, which require the deployment of complicated RSVP protocol and LDP protocol. The current technical evolution trend of the network towards SRv6 is to build a simplified network for unicast without deploying protocols or signaling such as LDP and RSVP. Traditional stateful multicast cannot support simplified networks like SRv6, which does not conform to the network development trend.

# 2. Technical Advantages of New Stateless BIER Multicast

As a new multicast forwarding technology, BIER encapsulates a BIER packet header for each multicast packet, and encapsulates BIER multicast receiver information in the BIER packet header. The BIER router forwards BIER multicast packets in accordance with the information in the BIER packet header, and does not maintain the multicast forwarding state information. BIER encapsulation isolates specific multicast services from the network layer. On the network, P routers no longer maintain the forwarding state of each multicast group for each VPN, and BIER routers do not perceive upper-layer multicast services, so that P routers are stateless to multicast. The stateless feature of BIER multicast eliminates the pressure on networks caused by large-scale deployment of multicast services.

BIER multicast uses the BIER extension of the traditional link state IGP protocol, such as ISIS or OSPF. The key BIER information such as BFR-id, SD and BSL are flooded to the network through BIER-prefix. Each BIER router generates a BIER forwarding table according to the key information. The BIER router forwards multicast packets according to the BIER forwarding table without querying the multicast forwarding table.

# 3. Basic Principle and Architecture of BIER

## 3.1. Basic Principle of BIER

Bit Index Explicit Replication (BIER) is a new multicast technology. Unlike the traditional PIM multicast protocol, the BIER protocol provides a stateless multicast forwarding mechanism. BIER determines the multicast receiver (BIER Egress) information at the BIER Ingress. The intermediate nodes do not need to maintain any multicast flow forwarding state information (Group, Ingress, Egress). BFIR is the BIER router that is closest to the multicast source. The local BIER forwarding table is generated by calculation based on the IGP BIER link state library, while the BIER link state library is generated by flooding based on the BIER extension of the IGP (ISIS/OSPF) protocol.

The basic principle of BIER is simple and efficient. Each BIER router can be assigned a non-repeated, unsigned integer, BFR-id, which uniquely identifies the BIER router. Each BIER router uses a specific prefix (BFR-prefix) carrying important information such as BFR-id, SD, BSL, encapsulation type and BIFT-ID to flood in IGP. BFR-Prefix is usually the host address of the local Loopback. Each BIER router calculates the BIER forwarding table of the optimal path to other BFR-id according to the IGP algorithm or BIER algorithm, which is similar to the process of generating the IPv4/IPv6 forwarding table. BIER multicast designs a bit string (BitString) of a specific length to represent a group of BIER routers. From the right most of the BitString, each bit corresponds to a BFR-id. For example, the binary string "101" with BitString length (BSL) of 3 indicates two BIER routers with BFR-id 1 and 3, and the binary string "011" indicates two BIER routers with BFR-id 1 and 2. The binary string "00011" with the BSL of 5 can also indicate two BIER routers with BFR-id 1 and 2. Different BSLs affect the load efficiency of the BIER packet header. The larger the BSL, the lower the efficiency, but the larger the number of BIER routers that can be expressed. At present, the RFC8279 standard requires that all BIER routers must support the BSL value of 256. A BIER router can support multiple different BSLs or use different BSLs in different networks.

Multiple Sub Domains (SDs) can be designed for a large BIER network based on the network topology or geographical location for simplified management. For example, nationwide operators can set up SD networks for the eastern region, the western region, the southern region and the northern region. By default, there can be only one SD network. The BSLs and BFR-ids in the SDs are independent of each other. In each SD, the concept of sub identify (SI) is introduced to express more BFR-ids with a shorter BSL. For example, the 256 BIER routers in the eastern region are represented by 4 different SIs and the BitString with BSL of 64. Each SI represents 64 BIER routers in a province in the eastern region. It no longer requires the BitString with BSL of 256 for identification. Using SIs in an SD can reduce the length of BSL, and the load efficiency of BIER packets is higher. However, each SI corresponds to one BIER forwarding table in the node.

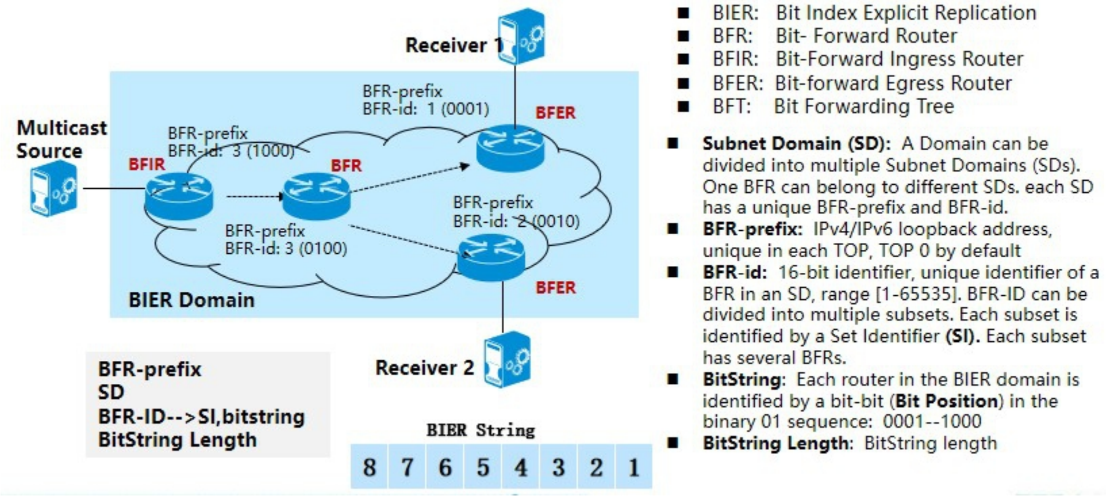


Figure 1 Basic Concepts of BIER

## 3.2. Three-Layer Architecture of BIER

### 3.2.1. Overlay Layer

IETF RFC8279 divides the BIER multicast architecture into three layers: Overlay, BIER and Underlay. The Overlay layer is responsible for information exchange on the multicast control plane, such as joining and leaving of user multicast flows between the BIER Egress node and the BIER Ingress node, and encapsulation, decapsulation and forwarding when multicast flows join and leave the BIER domain. The Overlay layer can be implemented through SDN, MP-BGP (MVPN), PIM, BMLD (BIER extension on MLD) and static configuration, among which MP-BGP and SDN are most common.

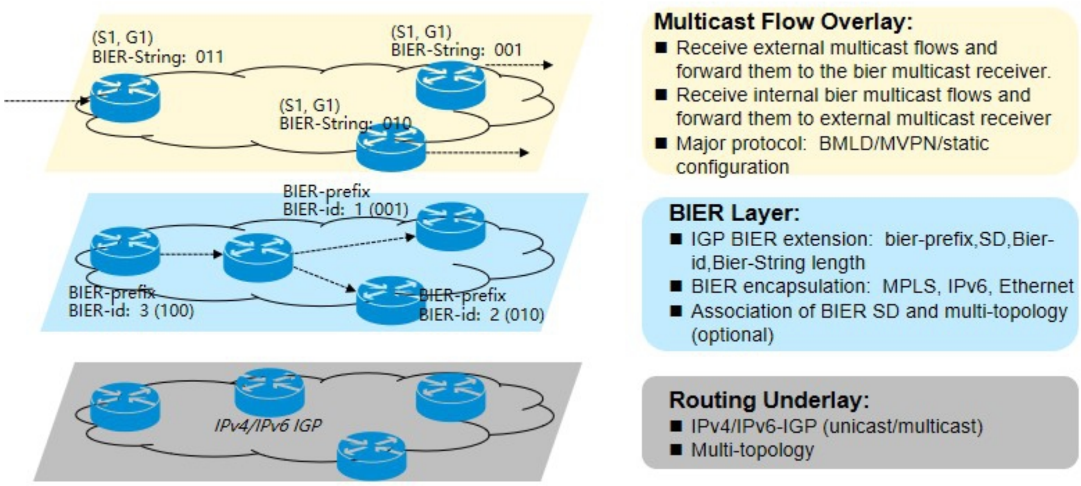


Figure 2 Three-Layer Architecture of BIER

### 3.2.2. BIER Layer

The BIER layer completes BIER routing information distribution, flooding, and calculates and local BIER forwarding table calculation and update. The BIER layer forwards BIER packets in accordance with the BIER forwarding table. Each node that forwards BIER packets decapsulates and re-encapsulates BIER packets. It decapsulates the BIER packet header, and obtains the key information carried by the BIER packet, such as BIFT-ID and BitString. The former is the index for the router to locate the BIER forwarding table, and the latter is the key for querying the BIER forwarding table, similar to using the destination IPv4 address as the key in IPv4 routing table query. The BIER node re-encapsulates the BIER packet header in accordance with the result of the BIER forwarding table and forwards the BIER packet. If this node is a multicast replication point, there are multiple query values. Each value indicates that the node will replicate and re-encapsulate a new BIER packet header and forward it.

A BIER router can have multiple BIER forwarding tables, and each BIER forwarding table has multiple entries. Each BIER forwarding table is associated with a BIFT-ID. The BIFT-ID is generated by coding or hashing of SD, SI and BSL. MPLS encapsulates BIER packets, and BIFT-ID is the MPLS label value of BIER. The BIER forwarding table entry is composed by a string of bit codes (Forwarding bit mask (F-BM) in RFC8296) and a neighbor node. Each string of F-BM indicates a set of BIER nodes reaching other nodes via this neighbor by the optimal path. The BIER node decapsulates the BitString in the BIER packet header, performs AND calculation using the BitString and the F-BM of each entry in the BIER forwarding table, and decides whether to copy the BIER packet to the neighbor node according to the calculation result. For the expression of F-BM, refer to Figure -11.

The BIER layer shields the perception of multicast services on the network layer. The intermediate BIER node does not perceive multicast services, does not establish traditional multicast distribution trees, or maintain the forwarding state information of each group. The BIER router only performs forwarding or replication according to the BitString of the received BIER packet and the local BIER forwarding table.

### 3.2.3. Underlay Layer

The Underlay layer is the traditional link state routing protocol layer. It uses the extended TLV attributes such as ISIS and OSPF link state protocols to carry the BIER information of this node. As a result, BIER inherits many features of ISIS and OSPF protocols, such as supporting FRR and load sharing. The BIER forwarding table convergence is synchronous with ISIS or OSPF convergence, with the speed reaching milliseconds. As shown in the following figure, the BIER extension of ISIS adds a BIER-SUB-TLV that carries important information such as BFR-id and SD. At the same time, multiple BIER-SUB-SUB-TLV are added to carry the encapsulation type (such as MPLS encapsulation, non-MPLS Ethernet encapsulation, and non-MPLS IPv6 encapsulation), BSL and the maximum SI value, label value or BIFT-id.

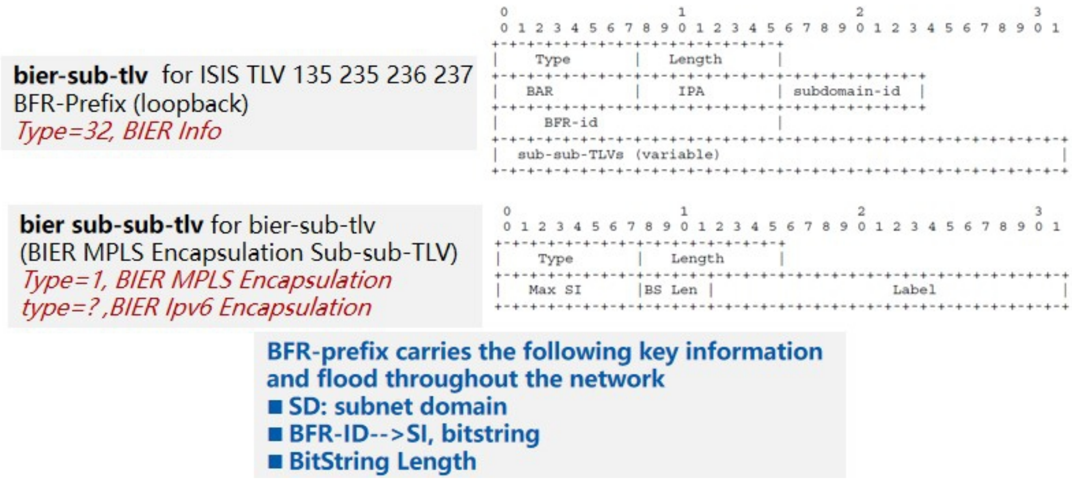


Figure 3 Extended TLV of BIER at Underlay Layer

# 4. BIER Packet Format and Encapsulation Type

## 4.1. BIER Packet Format

IETF defines three types of BIER packets: MPLS encapsulation, non-MPLS Ethernet encapsulation and non-MPLS IPv6 encapsulation to meet different networking requirements. Different BIER encapsulation types have the same BIER packet header, as shown in the dark blue part in the figure below. A multicast packet is encapsulated with a BIER packet header when entering the Ingress node of BIER. The BIER packet header is decapsulated to restore the multicast packet when the multicast packet leaves the BIER Egress node.

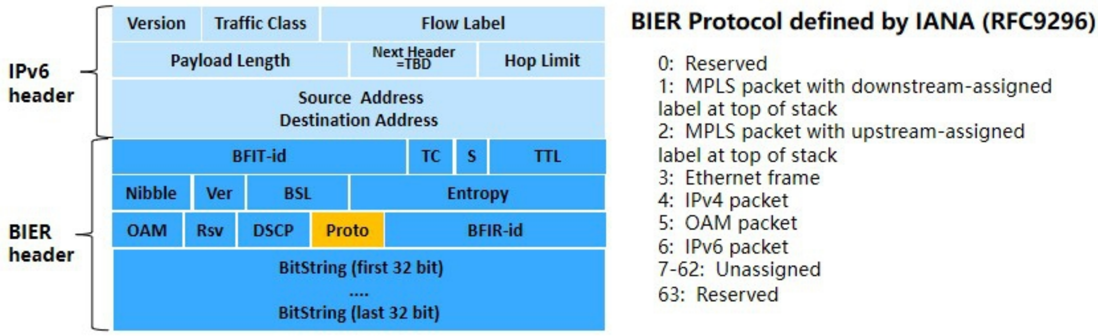


Figure 4 BIER Packet Header Format

BIFT-id: BITF used for packet forwarding, Label value in MPLS encapsulation. Non-MPLS encapsulation (Ethernet or IPv6) uses (SD, SI, BSL) for mapping or coding.

TC: Traffic class, same as TC in MPLS encapsulation, refer to RFC5462.

S: Label stack bottom label, same as S bit in MPLS encapsulation. Refer to RFC3032.

TTL: Same as TTL in MPLS encapsulation. Refer to RFC3032.

Nibble: Fixed value 0101, used to differentiate BIER encapsulation and MPLS ECMP function.

Ver: Version number. The current value is 0, referring to the version in experiment.

BSL: Length of BitString (log2(k)-5)m, used for offline analysis.

Entropy: Supports ECMP. For the same Entropy+BitString, select the same path.

OAM: The default value is 0. Ping/trace can be used, forwarding and QoS not affected.

RSV: Reserved bit. The default value is 0 is not used.

DSCP: Not used in MPLS encapsulation. Used in non-MPLS encapsulation.

Proto: Payload packet type, standardized by RFC.

BFIR-id: BFR-ID value of the first BIER router when the multicast packet enters the BIER domain.

BitStirng: Represents a set of BFER routers with SD and SI.

Based on the Protocol field in the BIER, the BIER load may be IPv6/IPv4 packets, MPLS or Ethernet packets. BIER supports both IPv4 and IPv6 multicast services. By means of upstream label allocation, it can also support multicast VPN services.

## 4.2. BIER Encapsulation Type

BIER supports MPLS encapsulation and non-MPLS encapsulation. Non-MPLS encapsulation supports Ethernet encapsulation and IPv6 encapsulation. At present, MPLS encapsulation and non-MPLS Ethernet encapsulation have been standardized, while IPv6 encapsulation of BIER has not been standardized yet. There are multiple BIER IPv6 encapsulation versions.

### 4.2.1. BIER Ethernet Encapsulation Format

IANA defines the BIER Ethernet encapsulation type 0XAB37, in which the BIER packet header is right behind the Ethernet header. The Ethernet encapsulation of the BIER is very simple and efficient, as shown in the figure below. The Ethernet type 0xAB37 indicates that its payload is BIER packets. The format of the BIER packet header is shown in Section 4.1. The Protocol field in the BIER packet header can further identify the content of the upper layer protocol. If the protocol field in the BIER packet is 2, it identifies the MPLS label allocated by the upstream and is usually used to implement the multicast VPN service. For detailed information of BIER Ethernet encapsulation, please refer to RFC8296 “Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks”。

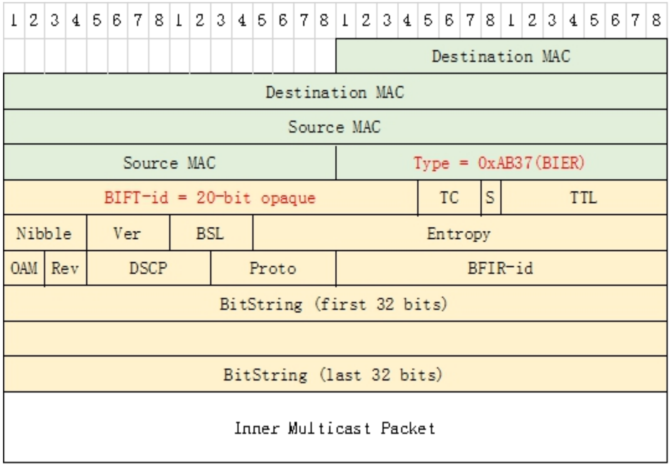


Figure 5 BIER Ethernet Encapsulation Format

## 4.2.2. BIER MPLS Encapsulation Format

The MPLS encapsulation of BIER uses the MPLS type defined by IANA: 0x8847. 0x8847 indicates that Ethernet packets are encapsulated MPLS packets, and the BIER packet type is determined by the label value range. Like the IPv4/IPv6 protocol, the label management module allocates an independent label range for BIER encapsulation. MPLS encapsulation reuses the first four bytes of the BIER packet header, where BIFT-ID, the first 20 bits of the BIER packet header, is the BIER label value allocated by the label management module. The MPLS encapsulation sequence of BIER is Ethernet, BIER packet headers, and upper-layer protocol. It should be noted there is no independent MPLS label, as shown in the figure below. After resolving the first four bytes (MPLS-BIER) in the BIER header, the MPLS module enters different protocols, including the BIER processing flow, and determines the corresponding forwarding table according to the label value. For detailed information of BIER MPLS encapsulation, please refer to RFC8296 “Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks”.

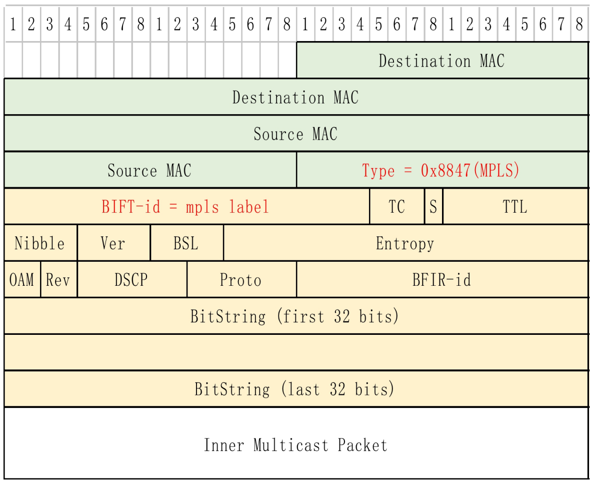


Figure 6 BIER MPLS Encapsulation Format

## 4.2.3. BIERin6 Encapsulation Format

Currently, ZTE supports BIERin6 encapsulation and draft-zhang-bier-bierin6-04. This draft recommends adding a new BIER protocol type to the Next Protocol of IPv6. The specific type value is to be allocated by IANA, which is represented by TBD in the following figure. The BIER packet header serves as IPv6 load. RFC82000 has defined the protocols supported by the IPv6, including IGMP, IPv4, TCP, IPv6, UDP, Ethernet and shim6. The upper layer protocol type supported by IPv6 and the extension header supported by IPv6 are implemented by using different values of the same field (Next Header). For example, if NextHeader is 0, it identifies the extension header of IPv6 hop-by-hop. If NextHeader is 43, it identifies the route extension header. SRv6 uses this extension header. The Protocol field in the BIERin6 encapsulation complies with the definition of RFC8296. The specific encapsulation format is as follows:

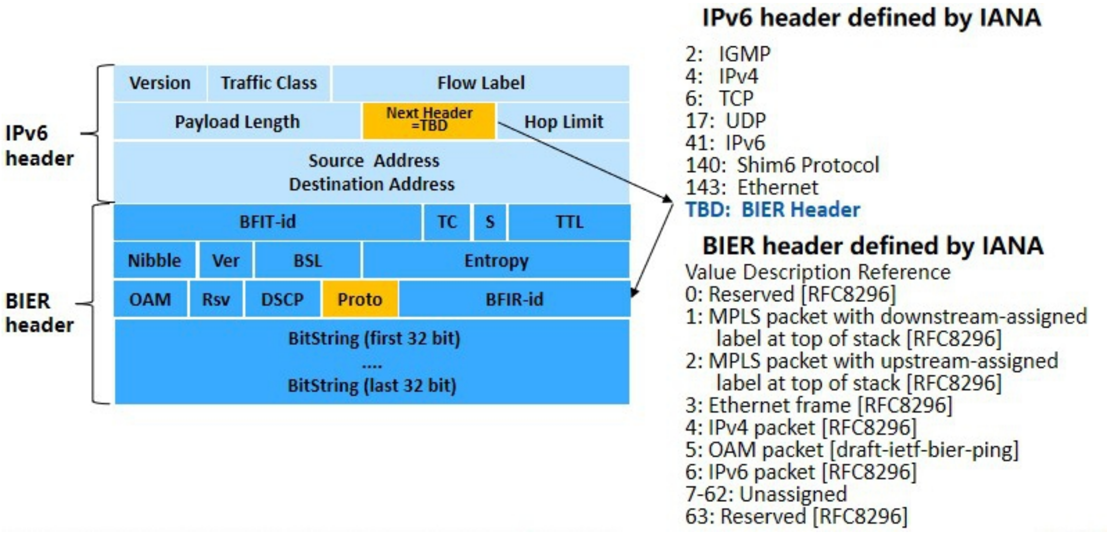


Figure 7 BIERin6 Encapsulation Format

The encapsulation of BIERin6 is seamlessly integrated with other IPv6 extension headers. BIER packets can be placed behind IPv6 extension headers such as Hop-by-Hop Options header and Destination Options header, and the load is identified as BIER packets by NextHead in the extension headers. In general, BIER packets are processed hop by hop. It is not recommended to fragment and encrypt or decrypt BIER packets in the BIER domain. The fragmentation, encryption and decryption of multicast service packets are processed at the upper layer, while the BFIR node only performs simple BIERin6 encapsulation for the multicast service.

# 5. ISIS Extensions Support BIER

## 5.1. ISIS Extensions sub-tlv and sub-sub-tlv Support BIER

Link state protocols such as ISIS and OSPF are extended to support BIER and are standardized (RFC8401 and RFC8444). At present, mainstream manufacturers such as ZTE/Huawei/Nokia have realized ISIS BIER extension, and OSPF BIER extension is also under development. The BIER-SUB-TLV and BIER-SUB-SUB-TLV of ISIS extension carry key BIER information such as BFR-id, BSL and SD. BIER-SUB-TLV is used together with ISIS TLV 135, 235, 236 and 237. It content is shown in the figure below.

**bier-sub-tlv**

```
 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     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    BAR        |     IPA       | subdomain-id  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      BFR-id                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  sub-sub-TLVs (variable)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8 ISIS -SUB-TLV

TYPE: BIER-SUB-TLV, the value is 32

Length: Changing value

BAR: BIER algorithm, used to calculate the path to BFER

IPA: IGP algorithm, enhance or improved IGP algorithm, can replace BAR algorithm.

Subdomain-id: Represents an SD

BFR-id: The 16 bits allocated to this route have no compliant integer.

Sub-sub-TLV: Optional sub-sub-TLV. The format is shown in the figure below.

A router can support multiple encapsulation modes such as MPLS encapsulation, Ethernet encapsulation and Bierin6 encapsulation. Different encapsulation modes are expressed by corresponding parameters. ISIS defines sub-sub-tlv to indicate different encapsulation types. A sub-tlv of a BIER can carry multiple different sub-sub-tlvs. The sub-sub-tlv in MPLS encapsulation that has been defined is as follows:

```
 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     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Max SI      |BS Len |                     Label             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9 ISIS -SUB-SUB-TLV (MPLS)

Type: The current value is 1, representing MPLS encapsulation

Length: Changeable value

Max SI: Maximum number of SIs it can support.

BS Len: Codes representing BSL, 4 bits.

Label: The first label in the label range

The sub-sub-tlv in Ethernet encapsulation is similar to the sub-sub-tlv in MPLS encapsulation. The major difference is that the type is 2 and that the 20-bit BIFT-id is used to replace the Label value in MPLS encapsulation, as shown in the figure below.

**BIER sub-sub-tlv (Ethernet encapsulation)**

```
 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     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Max SI     |BS Len |                 BIFT-id               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10 ISIS -SUB-SUB-TLV (Ethernet)

The BIFT-id of each SI is BIFT-id (initial value)+SI (value). If the BIFT-id (initial value)+SI value exceeds 20 bits, an error is reported.

## 5.2. ISIS BIER Routing Table and Forwarding Table

The sub-tlv and sub-sub-tlv in ISIS extensions carry the key BIER information: SD, BSL and BFR-ID, and flood in the ISIS network through BIER-prefix, including leading to the ISIS-L1 router. The BIER router uses the IGP algorithm or BIER algorithm to generate the route to the BFR-prefix, that is, the route to each BFR-id. Each BFR-prefix carries a sub-tlv corresponding to the information of one BFR-id. As shown in the following figure, R4 calculates the BFR-prefix routing table (BIRT) to R1, R2 and R3, that is, the routing table corresponding to the BFR-id. R4 combines several BFR-id (0: 0001/0010/0100) into one F-BM (Forward Bit-Mask) (0:0111) to generate the BFR forwarding table (BIFT)

The BIER router can calculate or allocate multiple BIFTs according to various factors such as protocol, topology and BSL. Each BIFT has a unique index BIFT-id, which corresponds to the BIFT-id field in the BIER packet header. The BIFT-id in the BIER packet is used as an index to reach the correct BIFT. The BIFT-id in the BIER packet is from the BIFT-id field in the sub-sub-tlv advertised by the neighbor router and is not generated locally. In MPLS BIER encapsulation, the BIFT-id is a 20-bit Label calculated based on SI and initial label. In Ethernet encapsulation, the BIFT-id is a 20-bit BIFT-id calculated based on SI and initial BIFT-id.

**R4 Bit Index Routing Table (BIRT)**

Figure -1

| BFR-prefix | BFR-ID | Neighbor |
|---|---|---|
| R1 | 0:0001 | R3 |
| R2 | 0:0010 | R3 |
| R3 | 0:0100 | R3 |

**R4 Bit Index Forwarding Table (BIFT)**

Figure -2

| Forward Bit-Mask(F-BM) | Neighbor |
|---|---|
| 0:0111 | R3 |

Figure -3

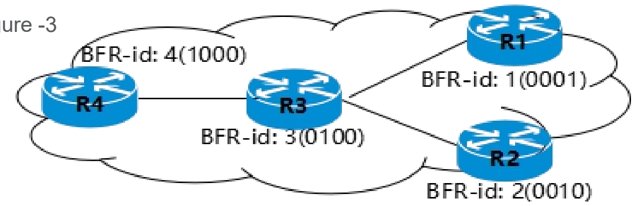


Figure 11　Generation of BIER Forwarding Table

# 6. BIER Forwarding Procedure

Multicast packets enter the BIER domain after the BIER packet header is encapsulated through the BFIR router. In the BIER domain, the BIER router locates the BIFT table according to the BIFT-id carried in the BIER packet header, and queries the BIFT according to the carried BitString for forwarding or replication. After arriving at the BFER node, the BIER packets are decapsulated into multicast packets, and forwarded after the multicast route is searched according to the multicast address. The BFIR determines the BitString in the BIER packet according to the Overlay layer information, that is, one multicast group collects all BFERs. The BFIR assembles the BIER packet header according to the BIFT-id advertised by the neighbor and the local BIER configuration information, and sends the BIER packet using the encapsulation type advertised by the neighbor, for example, MPLS encapsulation, Ethernet encapsulation or IPv6 encapsulation. The BIER forwarding procedure is described below with a specific example.

There are six BFR routers (A, B, C, D, E and F) in the BIER domain. D, F, E and A act as BFIR or BFER, with BFR-id = 1, 2, 3 and 4. The length of the BitString in the BIER domain is 4, and the SI is 0. The BFR-ids corresponding to the four routers (D, F, E and A) are 0:0001, 0:0010, 0:0100 and 0:1000 respectively. Other BFR routers are not allocated with BFR-ids, as shown in the figure below. Each router calculates and generates BIFT to other BFR-id according to IGP or BIER. Router A serves as the BFIR node for 239.1.1.1 multicast, while nodes D and E serve as BFER to apply for joining the multicast group 239.1.1.1. The forwarding procedure of BIER packets in the BIER domain is as follows:

- The BFIR (router A) receives the multicast packets of the 239.1.1.1, and searches the ingress interface and egress interface information of the multicast route. The egress interface is directed to an index of the BFER-List. The BFIR converts BFER-List into BitString, and confirms BIFT-id corresponding to multicast according to SD, SI and BSL. The BFIR uses the BitString to query the local BIFT table to obtain the key information of the corresponding egress interface for BIER forwarding and replication, such as the encapsulation type and BIFT-id advertised by the neighbor. According to the BFER-list, the BFIR deduces that the BitString of the encapsulated multicast packet is 0101, corresponding to nodes D and E (overlay message confirmation). At the same time the BFIR uses the BIFT-id advertised by neighbor node B to encapsulate the multicast packet. The BFIR sends the multicast packet to Router B after encapsulation.
- After node B receives the BIER packet sent from node A, it resolves the BIER header to obtain BIFT-id and BitString, searches for the specified BIFT forwarding table according to the BIFT-id, uses BitString (0101) and BIFT to search the table to obtain each egress neighbor and the corresponding BitString, and obtains the BIER encapsulation type according to the egress neighbor.

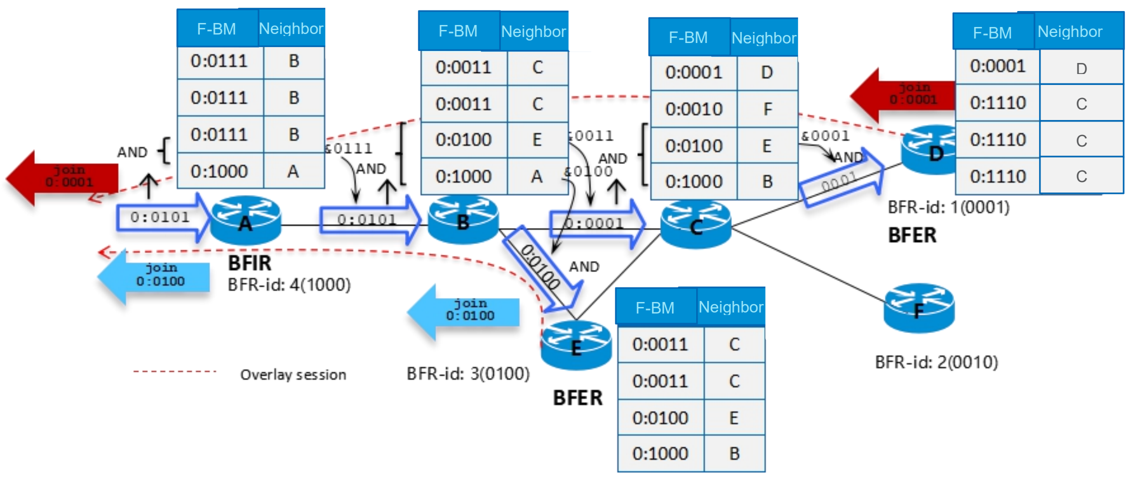


Figure 12 BIER Forwarding Procedure

- After receiving the BIER packet from Node B, Node C processes the packet in the same way as Node B, and sends a BIER encapsulation packet to Router D. The BIER packet header contains such information as BIFT-id advertised by Node D with BitString being 0001, and is sent using the BIER encapsulation type supported by Node D.
- After receiving a BIER packet from Node B, BFR node E processes the packet in the same way as node B. Because the AND calculation value of BitString (0100) and the F-BM in the third entry in BIFT is 0100, and the neighbor corresponding to this entry is E, it is confirmed that Node E is BFER. According to the query result of BIFT, Node E decapsulates the BIER packet, and searches the multicast routing table for forwarding according to the multicast address.
- After receiving a BIER packet from node C, node D processes the packet in the same way as node E. It is confirmed that Node D is BFER. It decapsulates the BIER packet, and searches the multicast routing table for forwarding according to the multicast address.

# 7. BIER Multicast Application Scenarios

## 7.1. Application of BIER in IPTV/OTT Scenario

IPTV is a point-to-multipoint application service in the network. It can provide live video streaming service and on-demand streaming program transmission service. For example, live streaming services of sports competitions and concerts belong to the former, while reality show programs whose progress can be controlled by on-demand playback belong to the latter. The IPTV live streaming service requires establishing a multicast distribution tree of the video source from the Egress node to the Ingress node, so as to save the backbone network bandwidth. Affected by the pandemic in 2020, new live streaming services are booming on the network, for example, various teaching platforms such as Tencent Class, Xueersi Class and various business video conferencing systems such as Zoom. The on-demand streaming programs can be pushed to the nearest edge CDN node through multi-level Content Delivery Networks (CDN). End users of on-demand streaming programs join the nearest edge CDN nodes through multicast. The distributed multi-level CDN network can save the bandwidth of the WAN and reduce the pressure of massive concurrent users on a single POP node and network. Large content providers and game operators can support massive concurrent users through distributed multi-level CDN networks. The distributed multi-level CDN network implements reliable content transmission and management through the application layer. Over The TOP (OTT) is not a common type of IPTV service. The major difference lies in that the video source of the OTT service is stored in another network. The OTT service needs to traverse two or more networks. For example, when watching Tencent videos through a mobile phone, the mobile phone terminal is in the operator's network while the video source is stored in Tencent's IDC equipment room. Tencent video traffic needs to traverse Tencent's IDC network and the operator's wired and wireless networks. The OTT service needs multicast service supporting traversal.

When BIER technology is used in the IPTV service, it is not necessary to create the multicast distribution tree from the Egress node to the Ingress node. The intermediate nodes do not run the multicast routing protocol or maintain the multicast forwarding state. In a typical IPTV network, the Egress node runs the IGMP/MDP protocol to obtain the channel joining or leaving message of the terminal. After that, the Egress node sends a channel joining or leaving message to the Ingress node. The Egress node also supports static configuration of joining or leaving a specific channel. The BIER processes the channel joining and leaving in a completely different way. The BIER directly maps the relationship between each channel and the receiver at the Ingress node. It is very convenient to add new channels in the BIER solution of IPTV. Adding new multicast mapping on the Ingress node can push all channels to all Egress nodes.

When the BIER technology is used in OTT services, the edge CDN node uses the powerful scaling feature of BIER to implement cross-domain multicast transmission. The BIER can use both the MP-BGP mode and the SDN mode to implement the cross-domain transmission. BIER also supports the static traversal mode, that is, the multicast traffic sent by the Egress node in one domain serves as the multicast source of the Ingress node in another domain. The ASBR in the CDN domain terminates the request from the OTT user in another domain, and sends the local cache video or the CDN video that is nearest to the ASBR to the OTT user. As a layer-4 application, the CDN can also implement simple applications across any different network boundaries. The CDN user in the remote BIER domain receives multicast packets as the source of the Ingress PE node in the domain to implement BIER forwarding in the domain.

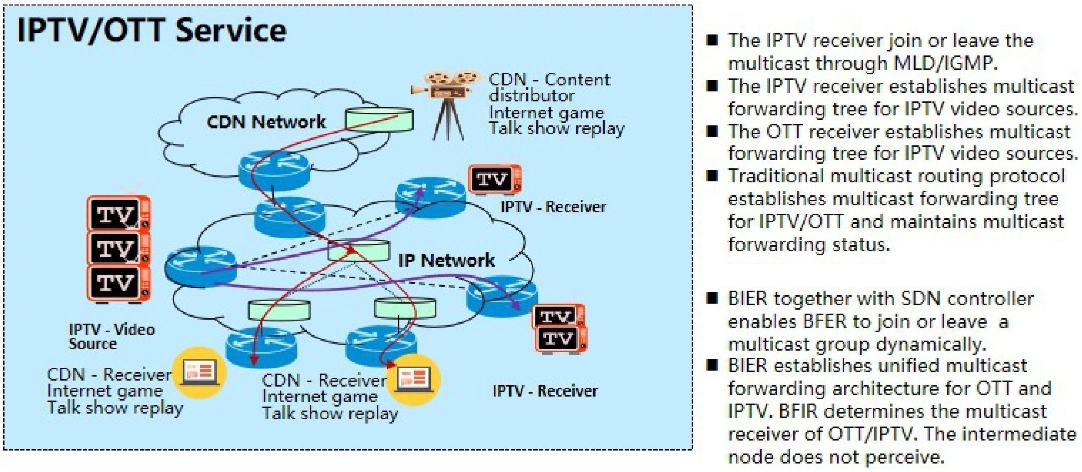


Figure 13 Application of BIER in IPTV/OTT Scenario

## 7.2. Application of BIER in Multicast VPN Scenario

The traditional multicast VPN needs the public network router to maintain the multicast VPN state. The operators cannot provide large-scale multicast service for all VPN customers because their network equipment cannot maintain the forwarding state of each multicast group for each VPN. The more the VPN customers and the number of groups, the more multicast state needs to be maintained by the operator equipment ($O(N^2)$). To reduce the number of multicast state maintained by the equipment, telecom operators have used head-end replication or multicast tree technologies, but do not completely solve the problems caused by substantial and complicated multicast states.

Unlike traditional multicast protocols such as PIM, BIER provides a stateless multicast forwarding mechanism. The first node of BIER multicast (BIER Ingress) determines the information of the multicast flow receiver (BIER Egress). The intermediate node forwards the multicast flows according to the BIER forwarding table and does not need to maintain the multicast state information. It is applicable to operators carrying out the multicast service of VPN. The Overlay layer of the BIER implements multicast VPN control information. The BIER Ingress node encapsulates BIER packets and VPN information according to the information of the multicast VPN receiver. The intermediate node implements BIER forwarding, without perceiving multicast or VPN. The BIER Egress node decapsulates the BIER information, and queries the corresponding VPN multicast routing table according to the inner VPN information for forwarding, as shown in the figure below.

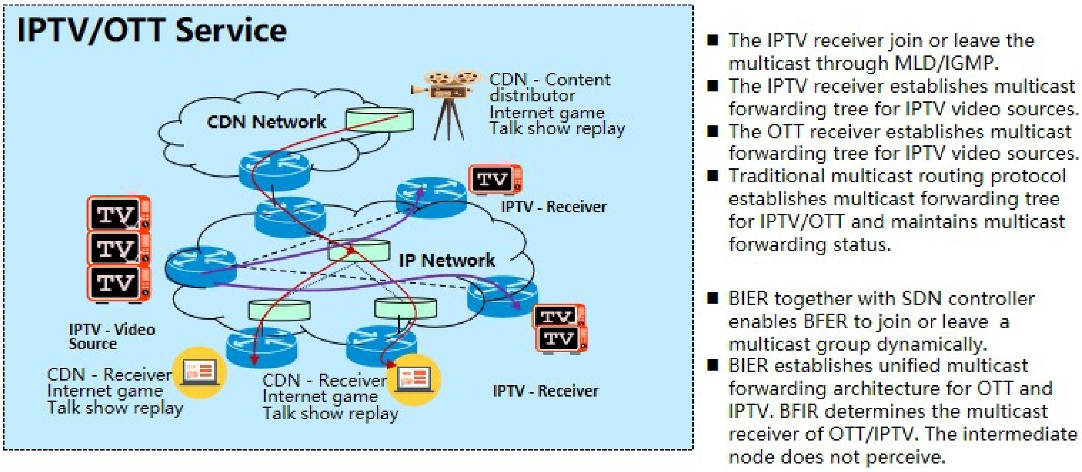


Figure 14 Application of BIER in VPN Multicast Scenario

## 7.3. Application of BIER in Financial Scenario

The multicast service is also widely used in financial services, most typically in stock service. The stock quotation data is transmitted to the operation divisions of securities companies in different places from the Shanghai/Shenzhen Stock Exchange through multicast service. These financial data need to arrive at each node safely with the optimal and determinate delay. No operation division of a securities company can stand to receive the stock quotation data from the exchange 100 milliseconds later than other securities companies.

Traditional multicast routing protocols such as PIM/mLDP/RSVP-TE do not completely solve the above problem. The current multicast trees are forwarding trees established from receivers to multicast sources, which is driven by multicast receivers. It is not necessarily the optimal path. When the number of multicast entries increases, the increasing number of multicast forwarding trees in the network will lead to changes in delay of multicast receivers. Some multicast receivers may receive multicast packets later than other multicast receivers. Complicated multicast protocols may affect data security.

BIER uses the existing unicast protocol to establish the optimal path from the source to the receiver. Even if the number of multicast packets increases, the characteristics of the optimal path from the source to the destination are not changed. Regardless of the number of multicast packets, the BIER convergence time is as fast as the unicast routing protocol. BIER always provides the optimal and determinate delay. BIER has remarkable advantages in assuring the fairness of financial multicast services compared with traditional multicast routing protocols.

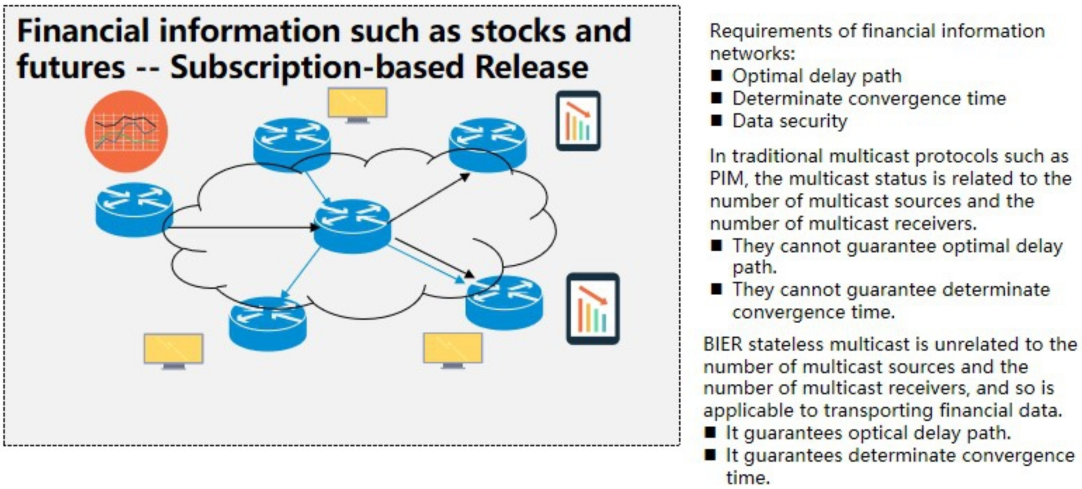


Figure 15 Application of BIER in Financial Scenario

## 7.4. Application of BIER in EVPN Scenario

EVPN L2VPN as a substitute for VPLS L2VPN is very popular now and has been widely used in actual deployment. EVPN-based L2VPN has such advantages as supporting AC dual-homing access, but it also has the requirement to process a large number of BUMs (Broadcast, Unknown unicast and Multicast). The BUM packets of one EVI instance need to be flooded to all PE routers that support this EVI instance. The first node replication enables EVPN to transmit BUM packets between a group of PE nodes. This is a common solution used in actual deployment. The first node replication copies the BUM packets of the Egress PE of the same EVI instance on the Ingress PE node. The number of the BUM packets copied on the Ingress PE node is equal to the number of the Egress PE nodes of the EVI instance. Therefore, Ingress PE is required to have both high multicast replication & forwarding performance and high bandwidth. The more the PE nodes in EVPN, the higher the performance and bandwidth requirements of the PE nodes.

The BIER implements BIER encapsulation and forwarding of BUM packets through the BIER P-tunnel function carried by the PMSI attribute of EVPN. BIER forwarding does not require P nodes run traditional protocols such as PIM/mLDP/P2MP/MP2MP to maintain multicast forwarding status. The P node forwards the BUM packets in BIER encapsulation according to the BIER forwarding table. The P node does not perceive the multicast service, operate the multicast protocol, or maintain the multicast forwarding status. Normally, through the I-PMSI attribute of EVPN, one BIER Subnet Domain can correspond to multiple EVI instances. The Ingress PE allocates different labels to different EVI instances and advertises them to the downstream Egress node. The Egress node decapsulates the BIER packet to identify the EVI label and forwards BUM packets to enter the corresponding EVI instances. If the Ingress PE uses the Inclusive-PMSI interface to send BUM packets of multiple different EVI instances, the BUM packets may reach the Egress PE node without a receiver, which wastes the bandwidth. The BIER packet carries BitString that can precisely locate the Egress PE node. If the Ingress PE performs BitString coding and BIER encapsulation for the Egress PE node that receives BUM packets, it can prevent the above situation. From this perspective, BIER EVPN multicast supports both Inclusive Tree and Selective Tree. Although the Subnet Domain of a BIER can support any number of EVI instances, an independent BIER Subnet Domain can be set up for each EVI instance if required.

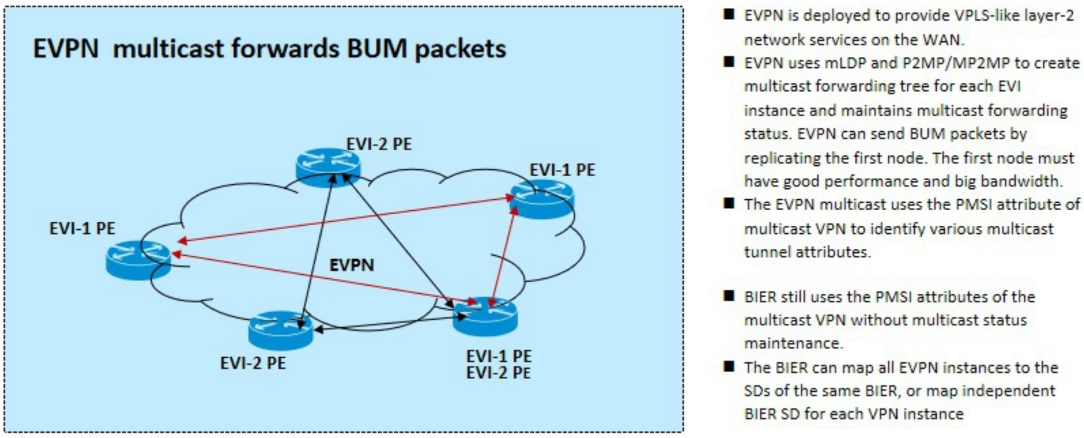


Figure 16 Application of BIER in EVPN Scenario

## 7.5. Application of BIER in Data Center Scenario

VXLAN is deployed in the data center on a large scale as a layer-2 transport network over a layer-3 network. The higher the capability of the data center supporting the multi-tenant and virtualization technologies, the higher the requirement for the number of VXLANs. A VXLAN maximally supports 16MB. Each VXLAN needs to forward L2 BUM packets between the VTP nodes in the data center. This requires the underlay network of the data center should support the multicast function. In the underlay network of the data center, the traditional multicast routing protocols such as PIM-SM can support the multicast requirement of small-scale VXLAN. When VXLAN reaches 16 M, the underlay network cannot support the 16M multicast forwarding tree. Even if the mapping between VXLAN and multicast is N:1, it may lead to non-optimized BUM forwarding path, resulting in unnecessary bandwidth waste. In general, data center administrators are not inclined to implement complicated multicast routing protocols inside data centers. The small-scale VXLAN multicast administrators implement multicast routing protocols through the first node replication. The first node replication can avoid complicated multicast protocols and simplify network protocol control, but has high requirements for network bandwidth and VTP node performance.

Deploying BIER in the data center can implement stateless forwarding of BUM packets without the need to create multicast forwarding tree or maintain multicast forwarding status. The mapping relationship between VXLAN and multicast is implemented on the first node or the SDN controller.

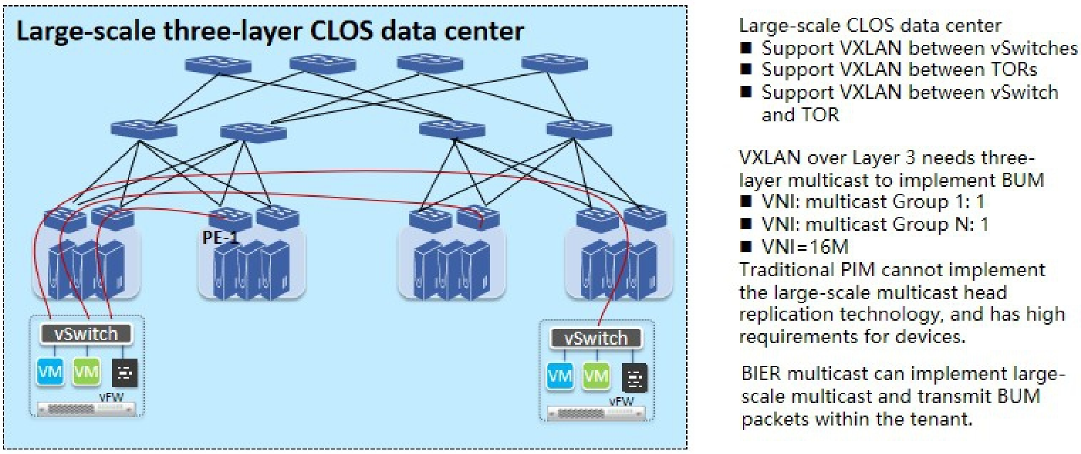


Figure 17 Application of BIER in Large-Scale Data Center Scenario

# 8. BIER-Related Standards

| No. | Classification | Standard | Description |
|---|---|---|---|
| 1 | Architecture | RFC8279 | BIER architecture, including concepts, terms, operating mechanisms and hierarchical models |
| 2 | Case | draft-ietf-bier-use-cases-11 | Description of BIER cases |
| 3 | Requirements | draft-ietf-bier-oam-requirements-09 | Description of BIER OAM requirements |
| 4 | Requirements | draft-ietf-bier-ipv6-requirements-04 | Description of BIER requirements in IPv6 scenarios |
| 5 | Data plane | RFC8296 | MPLS and non-MPLS network BIER encapsulation, including the structure of BIER Header and definition of the fields |
| 6 | BIER control plane | RFC8401 | ISIS extension for BIER-MPLS encapsulation, including related sub-TLV definition and protocol interactive processing mechanism |
| 7 | | RFC8444 | OSPFv2 extension for BIER-MPLS encapsulation, including related sub-TLV definition and protocol interactive processing mechanism |
| 8 | | draft-ietf-bier-lsr-ethernet-extensions-01 | IGP extension for BIER-Ethernet encapsulation, including related sub-TLV definition and protocol interactive processing mechanism |
| 9 | BIER control plane | draft-ietf-bier-ospfv3-extensions-01 | OSPFv3 extension for BIER, including related sub-TLV definition and protocol interactive processing mechanism |
| 10 | | draft-ietf-bier-bar-ipa-06 | IGP extension for BIER path algorithm and constraint conditions |
| 11 | | draft-ietf-bier-idr-extensions-07 | BGP extension for BIER |
| 12 | | draft-zhang-bier-bierin6-04 | IGP extension for BIER in IPv6 channel |
| 13 | | draft-zwzw-bier-prefix-redistribute-05 | IGP extension for BIER prefix inter-domain re-distribution |
| 14 | Overlay control plane | RFC8556 | BGP extension for BIER carrying L3VPN and IP global multicast |
| 15 | | draft-ietf-bier-mld-04 | MLD extension for BIER Overlay |
| 16 | | draft-ietf-bier-pim-signaling-08 | PIM extension for BIER Overlay |
| 17 | | draft-ietf-bier-evpn-03 | BGP extension for BIER carrying EVPN BUM multicast |

| | | | |
|---|---|---|---|
| 18 | OAM | draft-ietf-bier-ping-06 | BIER ping and trace message format and processing mechanism |
| 19 | | draft-ietf-bier-path-mtu-discovery-07 | BIER path mtu discovery mechanism |
| 20 | | draft-ietf-bier-pmmm-oam-07 | BIER coloring performance measurement method |
| 21 | | draft-hu-bier-bfd-05 | BIER P2MP BFD message format and processing mechanism |
| 22 | Northbound interface | draft-ietf-bier-bgp-ls-bier-ext-06 | BGP-LS extension to support BIER topology information report |
| 23 | | draft-ietf-bier-bier-yang-06 | BIER-related YANG model |

# 9. Abbreviations

| Abbreviation | Full name | Description |
|---|---|---|
| **BIER** | Bit Index Explicit Replication | Explicit multicast traffic based on bit index |
| **BFR** | Bit-Forward Router | Router supporting BIER forwarding |
| **BFIR** | Bit-Forward Ingress Router | Edge router in BIER domain, BFR router connected to the multicast source |
| **BFER** | Bit-forward Egress Router | Edge router in BIER domain, BFR router connected to the multicast receiver |
| **SD** | Sub Domain | One BIER domain can be designed with multiple Subnet Domains (SD) corresponding to different topology, and one BFR can cross Subnet Domains |
| **BFR-prefix** | Bit-Forward Router prefix | IPv4/IPv6 loopback address of BFR. BIER information is advertised according through this interface |
| **BFR-id** | Bit-Forward Router Identifier | A 16-bit integer, uniquely identifying the BFR within the SD, range: [1-65535]. |
| **SI** | Set Identifier | BFR-id can be divided into multiple sub-sets, identified by Set Identifier (SI). There are several BFRs in the sub-set. |
| **BitString** | BitString | A binary strings representing the BFR in SD. Each Bit corresponds to one BFR in SD. |

| | | |
|---|---|---|
| **BP** | Bit Position | Specific Bit Position of BFR-id in BitString. |
| **BSL** | BitString Length | BitString length |
| **BFT** | Bit Forwarding Tree | Forwarding path tree of multicast flow in BIER domain |
| **BIRT** | Bit Index Routing Table | Routing table calculated using BRF-prefix by BIER algorithm |
| **BIFT** | Bit Index Routing Forwarding Table | BFR forwarding table organized and optimized according to BIRT and neighbor nodes. |
| **BIFT-id** | | Specific BIFT ID |
| **BM** | Bit-Mask | Forwarding Bit Mask in BIFT, Bit string aggregated by multiple different BPs |