PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 22, 2022

H. Li
A. Wang
China Telecom
H. Chen
Futurewei
R. Chen
ZTE Corporation

October 19, 2021

PCE based BIER Procedures and Protocol Extensions
draft-li-pce-based-bier-02

Abstract

This document describes extensions to Path Computation Element (PCE) communication Protocol (PCEP) for supporting the PCE based Bit Index Explicit Replication (BIER) deployment.

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 April 22, 2022.

Table of Contents

1. Introduction

[RFC8279] defines a Bit Index Explicit Replication (BIER) architecture where all intended multicast receivers are encoded as a bitmask in the multicast packet header within different encapsulations such as described in [RFC8296] . A router that receives such a packet will forward the packet based on the bit

position in the packet header towards the receiver(s) following a precomputed tree for each of the bits in the packet. Each receiver is represented by a unique bit in the bitmask.

Currently, multicast management information is mainly signaled by PIM [RFC2362] or BGP [RFC6514], which have some limitations in the deployment and process.

[RFC4655] defines a stateful PCE to be one in which the PCE maintains "strict synchronization between the PCE and not only the network states (in term of topology and resource information), but also the set of computed paths and reserved resources in use in the network." [RFC8231] specifies a set of extensions to PCEP to support state synchronization between PCCs and PCEs.

This document specifies PCEP protocol extensions to optimize the implementation of multicast source registration or revocation, receiver automatic discovery, and forwarding control of multicast data by using PCEP messages to transmit multicast management signaling, combining with the forwarding characteristics of BIER.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Terminology

The following terms are used in this document:

- BFR-id: BFR Identifier. It is a number in the range [1,65535]
- BGP: Border Gateway Protocol
- BIER: Bit Index Explicit Replication
- BIFT: Bit Index Forwarding Table
- FI: Forwarding indication
- IGMP: Internet Group Management Protocol
- IGP: Interior Gateway Protocols
- MLD: Multicast Listener Discover

- MRI: Multicast Receiver Information
- MSR: Multicast Source Registration
- PCC: Path Computation Client
- PCE: Path Computation Element
- PCEP: PCE communication Protocol
- PIM: Protocol Independent Multicast

4. Overview of PCE based BIER solution

PCE based BIER includes multicast source registration information management, multicast receiver information management and multicast data forwarding control.

Multicast source registration information includes registration and processing of multicast source information.

Multicast receiver information includes requesting multicast group, multicast source and BitPosition information of receiver-side PCC.

Multicast data forwarding control includes BitString processing and data forwarding.

PCRpt message and PCUpd message, described in [RFC8231], are used in the PCE based BIER processing.

This document specifies PCEP protocol extensions for multicast group management, including Multicast Source Registration (MSR) object, Multicast Receiver Information (MRI) object, Forwarding Indication (FI) object and Multicast Receiver Status(MRS) object.

4.1. Example of PCE based BIER Topology

An example of PCE based BIER topology for a BIER domain with a controller as PCE is shown in Figure 1. In this domain, node R1 and R7 are Bit-Forwarding Ingress Router (BFIR) and Bit-Forwarding Egress Router (BFER), respectively.

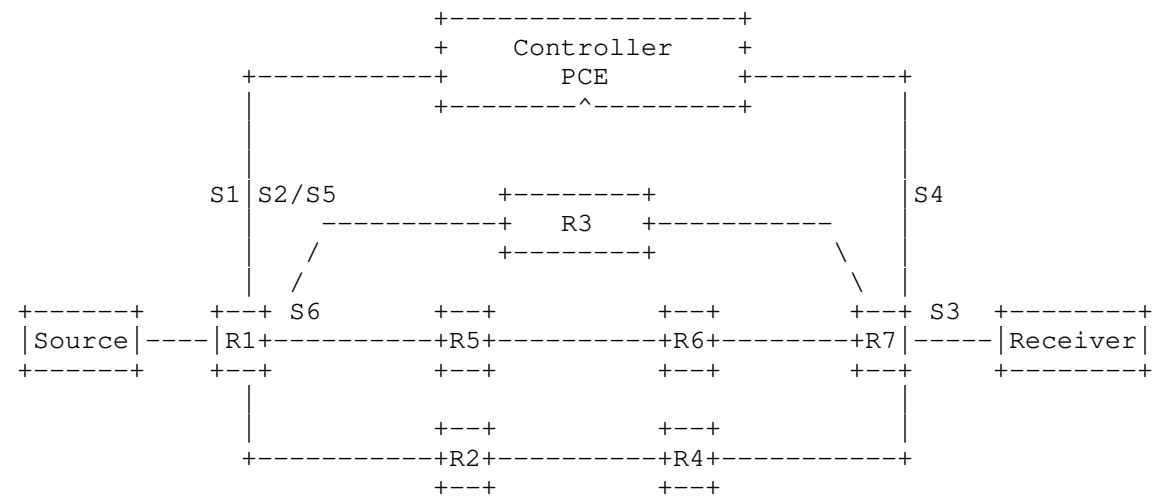

Figure 1: Example of PCE based BIER Topology(controller as PCE)

4.2. Basic Procedures

Step 1(S1): R1 sends multicast source information and authentication information to the controller about multicast information registration via PCRpt message.

Step 2(S2): The controller sends PCUpd message to R1, carrying authentication result.

Step 3(S3): Receivers send IGMP or MLD messages to R7 requesting to join or leave a multicast group.

Step 4(S4): R7 converts the IGMP or MLD messages into PCRpt message and sends it to the controller.

Step 5(S5): If the multicast group and multicast source information requested by the receiver has registered, the controller will send PCUpd message to R1 to start or stop forwarding, carrying BitString.

Step 6(S6): If R1 is ready to start forwarding, it will encapsulate BIER header and forward them based on BIFT and BitString when receiving multicast packets.

5. Capability Advertisement

During the PCEP initialization phase, PCEP speakers advertise stateful capability via the STATEFUL-PCE-CAPABILITY TLV in the OPEN

object. Various flags are defined for the STATEFUL-PCE-CAPABILITY TLV defined in [RFC8231] and updated in[RFC8232] and [RFC8281].

A new flag is added in this document, whose code point is TBD1:

B (BIER-MULTICAST-CAPABILITY, 1 bit): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker supports the capability of these new flag as specified in this document.

If a PCEP speaker receivers PECP message with the newly defined object, but without the B bit set in STATEFUL-PCE-CAPABILITY TLV in the OPEN object, it MUST:

- Send a PCErr message with Error-Type=10(Reception of an invalid object) and Error-Value TBD2(BIER-MULTICAST-CAPABILITY bit is not set).
- Terminate the PCEP session.

6. PCEP message

6.1. PCRpt message

MSR objectSection 7.1 should be included in the PCRpt message when PCC registers multicast source information with PCE.

MRI objectSection 7.2 should be included in the PCRpt message when PCC sends multcast join messages to PCE.

MRS objectSection 7.4 should be included in the PCRpt message when PCC inform PCE of the number of receivers.

The definition of the PCRpt message from [RFC8231] is extended to optionally include MSR object, MRI object and MRS object after the path object. The encoding from [RFC8231] will become:

```
   <PCRpt Message> ::= <Common Header>
                       <state-report-list>
Where:

   <state-report-list> ::= <state-report>[<state-report-list>]

   <state-report> ::= [<SRP>]
                      <LSP>
                      <path>
                      [<MSR>]
                      [<MRI>]
                      [<MRS>]
```

Where:
<path> is as per [RFC8231] and the LSP and SRP object are also defined in [RFC8231].

6.2. PCUpd message

MSR objectSection 7.1 should be included in the PCUpd message when PCE responds to the registration request.

FI objectSection 7.3 should be included in the PCUpd message when PCE sends the BitString to PCC to indicate the path of multicast data packets forwarding for PCC.

MRS objectSection 7.4 should be included in the PCUpd message when PCE inform PCC of the number of receivers.

The definition of the PCUpd message from [RFC8231] is extended to optionally include MSR object, FI object and MRS object after the path object. The encoding from [RFC8231] will become:

```
   <PCUpd Message> ::= <Common Header>
                       <update-request-list>
Where:

   <update-request-list> ::= <update-request>[<update-request-list>]

   <update-request> ::= <SRP>
                        <LSP>
                        <path>
                        [<MSR>]
                        [<FI>]
                        [<MRS>]
```

Where:
<path> is as per [RFC8231] and the LSP and SRP object are also defined in [RFC8231].

7. Object formats

7.1. Multicast Source Registration Object

The MSR object is optional and specifies multicast source information in multicast registration information management. The MSR object should be carried within a PCRpt message sent by PCC to PCE for registration. The MSR object should be carried within a PCUpd message sent by PCE to PCC in response to registration.

MSR Object-Class is TBD3. MSR Object-Type is 1.

The format of the MSR object body is:

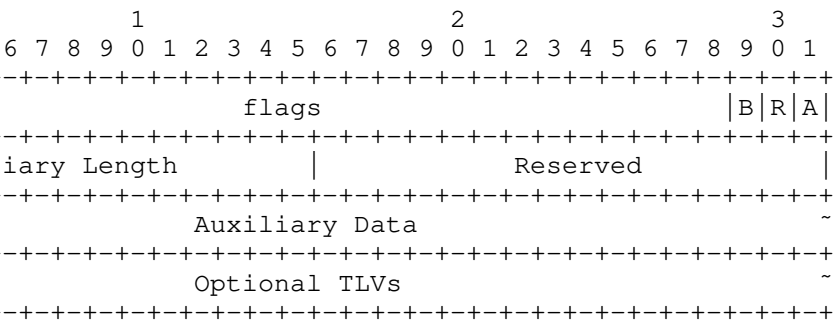

Figure 2: MSR Object Body Format

B(BIER multicast flag, 1 bit): The R flag set to 1 indicates that multicast protocol is BIER. The R flag set to 0 indicates that multicast protocol is not BIER.

R (Register flag, 1 bit): The R flag set to 1 indicates that the PCC is registering multicast information to the PCE. The R flag set to 0 indicates that the PCC revokes the register.

A (Authentication flag, 1 bit): The A flag set to 1 indicates success of registration. The A flag set to 0 indicates failure of registration or cancellation of registration. R and A cannot both be set to 0 or 1 in PCRpt message.

Auxiliary Length(8 bits): indicates the length of Auxiliary Data.

Auxiliary Data(Variable length): contains functional data such as authentication information.

MSR object could include three types of TLVs, namely Multicast Source Address TLV, BIER Information TLV, VPN Information TLV, as defined follows:

7.1.1. Multicast Source Address TLV

The format of the Multicast Source Address TLV is:

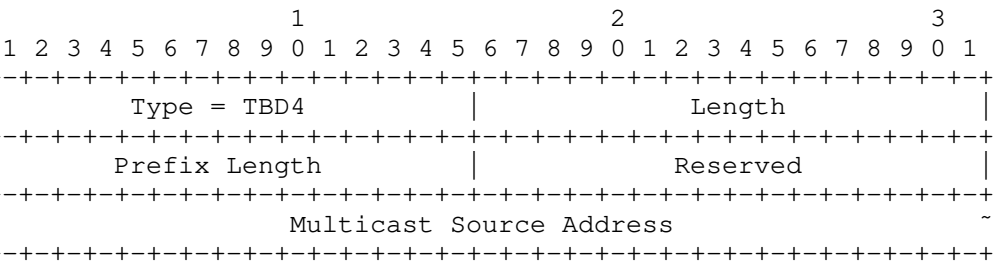

Figure 3: Multicast Source Address TLV Format

Type(16 bits): TBD4 is to be assigned by IANA.

Length: Variable.

Prefix Length(16 bits): indicates the length of multicast source address.

Multicast Source Address(Variable length): contains IPv4 or IPv6 address of the multicast source.

7.1.2. BIER Information TLV

BIER Information TLV is used to report router location information in the BIER domain. When the multicast flag in MSR, MRI, FI objects is set, BIER Information TLV should be included. The format of the BIER Information TLV is:

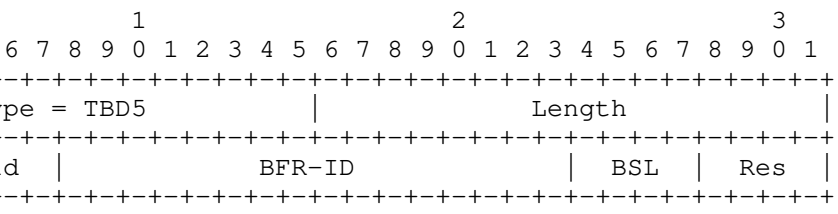

Figure 4: BIER Information TLV Format

Type(16 bits): TBD5 is to be assigned by IANA.

Length: Variable.

Subdomain-id(8 bits): Unique value identifying the BIER subdomain.

BFR-ID (16 bits): Identification of BFR in a subdomain.

BSL(BitString Length, 4 bits): encodes the length in bits of the BitString as per[RFC8296] , the maximum length of the BitString is 7, it indicates the length of BitString is 4096. It is used to refer to the number of bits in the BitString.

7.1.3. VPN Information TLV

VPN Information TLV is used to report VPN information about multicast sources and receivers. When the multicast flag in MSR, MRI, FI objects is set, VPN Information TLV should be included. The format of the VPN Information TLV is:

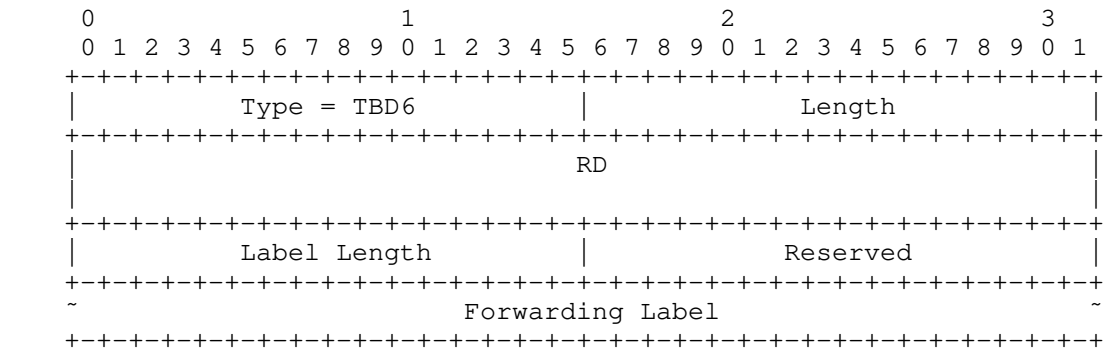

Figure 5: VPN Information TLV Format

Type(16 bits): TBD6 is to be assigned by IANA.

Length: Variable.

RD(Route Distinguisher, 8 bytes): indicates the VPN which the receiver used.

Label Length(16 bits): indicates the length of forwarding label Data, the length should be 0 ,32 bits or 128 bits.

Forwarding Label(Variable Length): contains MPLS label with 32 bit or IPv6 Segment Identifier with 128 bits.

7.2. Multicast Receiver Information Object

The MRI object is optional and specifies receivers' information for matching the multicast registration information. The MRI object should be carried within a PCRpt message sent by PCC to PCE in muticast joining or leaving.

MRI Object-Class is TBD7. MRI Object-Type is 1.

The format of the MRI object body is:

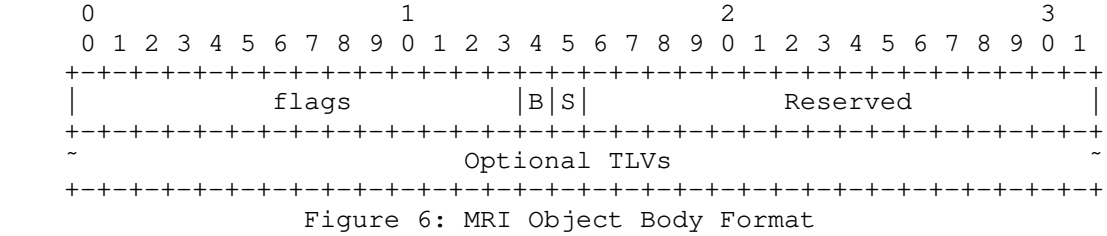

Figure 6: MRI Object Body Format

B(BIER multicast flag, 1 bit): The R flag set to 1 indicates that multicast protocol is BIER. The R flag set to 0 indicates that multicast protocol is not BIER.

S(Subscribe flag, 1 bit): The S flag set to 1 indicates that PCC delivers the message requesting to join PCE. The S flag set to 0 indicates that PCC delivers the message requesting to leave to PCE.

MRI object could include four types of TLVs, namely Multicast Source Address TLVSection 7.1.1, BIER INFO TLVSection 7.1.2, VPN Information TLVSection 7.1.3 and Multicast Group Address TLV. Multicast Group Address TLV is defined as follows:

7.2.1. Multicast Group Address TLV

The format of the Multicast Group Address TLV is:

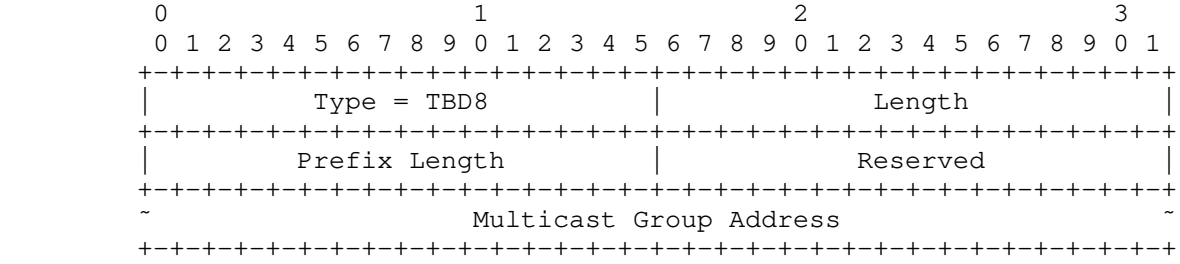

Figure 7: Multicast Group Address TLV Format

Type(16 bits): TBD8 is to be assigned by IANA.

Length: Variable.

Prefix Length(16 bits): indicates the length of multicast group address.

Multicast Group Address(Variable length): contains IPv4 or IPv6 address of the multicast group.

7.3. Forwarding Indication Object

The FI object is optional and used to indicate to the headend how to forward multicast data packets in the form of BitString. The FI object should be carried within a PCUpd message sent by PCE to PCC in multicast scenarios.

FI Object-Class is TBD9. FI Object-Type is 1.

The format of the FI object body is:

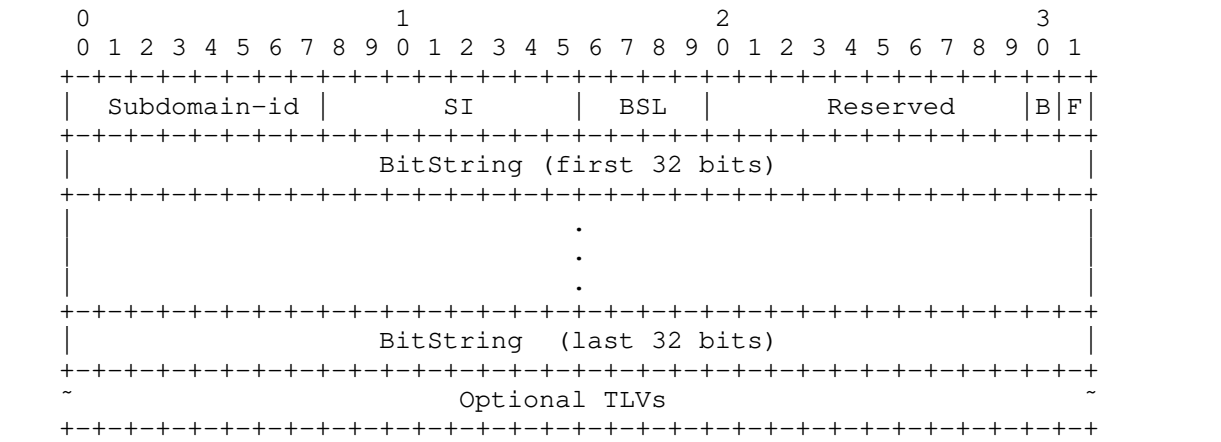

Figure 8: FI Object Body Format

Subdomain-id(8 bits): Unique value identifying the BIER subdomain.

SI (Set Identifier, 8 bits): encoding the Set Identifier used in the encapsulation for this BIER subdomain for this BitString length..

BSL(BitString Length, 4 bits): encodes the length in bits of the BitString as per[RFC8296] , the maximum length of the BitString is 7, it indicates the length of BitString is 4096. It is used to refer to the number of bits in the BitString.

B(BIER multicast flag, 1 bit): The R flag set to 1 indicates that multicast protocol is BIER. The R flag set to 0 indicates that multicast protocol is not BIER.

F(Forwarding flag, 1 bit): The F flag set to 1 indicates that the router may start forwarding multicast packets. The F flag set to 0 indicates that the router should stop forwarding multicast packets.

BitString(Variable length): indicates the path of multicast data packets forwarding for headend.

FI object should include three types of TLVs, namely Multicast Source Address TLVSection 7.1.1, VPN Information TLVSection 7.1.3 and Multicast Group Address TLVSection 7.2.1.

7.4. Multicast Receiver Status Object

The MRS object is optional and used to inform PCE of the number of receivers. The MRS object should be carried within a PCRpt or a PCUpd message for synchronize receiver information periodically, or PCRpt message for the leaving of receivers.

MRS Object-Class is TBD10. MRS Object-Type is 1.

The format of the MRS object body is:

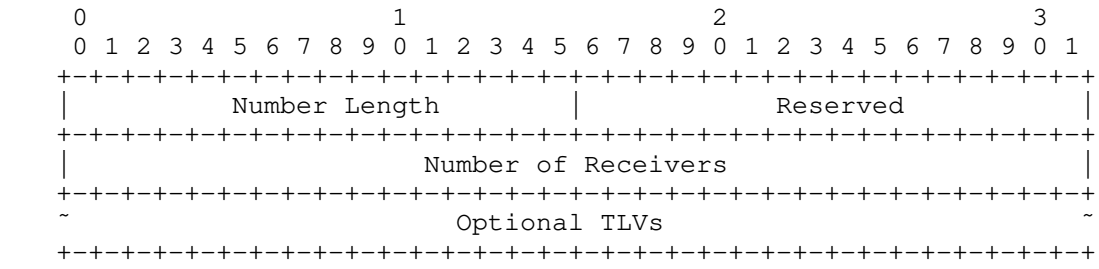

Figure 9: MRS Object Body Format

Number Length(16 bits): indicates the length of receiver number.

Number of Receivers(32 bits): indicates the number of receivers for a particular (S,G) tuple.

MRS object should include two types of TLVs, namely Multicast Source Address TLVSection 7.1.1 and Multicast Group Address TLVSection 7.2.1.

8. Procedures

8.1. Multicast source registration and revocation

For PCC-Registered multicast source, an ingress node sends a PCRpt message with MSR object to a stateful PCE, where R flag is set and A flag is not set. The registered authentication information can be passed through auxiliary data in MSR object.

Upon receiving the registration via PCRpt message, the stateful PCE MUST match local authentication rules based on the multicast information and auxiliary data in PCRpt message. If authenticated successfully, the PCE stores the multicast registration information into the database. In response, PCE MUST send a PCUpd message with MSR object to ingress node, where R flag is set. A flag is set only if authentication is successful.

For PCC-revoked multicast source registration, an ingress node sends a PCRpt message with MSR object to a stateful PCE, where R flag is not set and A flag is set.

Upon receiving the revocation via PCRpt message, in response, PCE MUST send a PCUpd message with MSR object to ingress node, where neither R nor A is set.

8.2. Joining and leaving of multicast receivers

When an egress node receives an IGMP or MLD message from a multicast receiver to join, the egress node should send a PCRpt message with MRI object to the PCE if no other receiver has sent the same request to it before.

If it is not the first time the PCE has received the same PCRpt message for join from the same egress node, this message should be ignored.

When an egress node receives an IGMP or MLD message from a multicast receiver to leave, the egress node should send a PCRpt message with MRI object and MRS object to the PCE if there are no other members in the requested multicast group. In MRS object, the number of receivers is zero.

8.3. BitString management

Upon receiving the join or leave request via PCRpt message, PCE needs to combine the BFR-id and SI of the egress node carried in PCRpt message with the BFR-id and SI of the ingress node and existed BitStrings in the database to create or update BitString. If there are members in the multicast group, the PCE should send a PCUpd message with FI object carrying the latest BitString to the ingress node, where F flag is set.

When receving multicast packets, the ingress node encapsulates BIER header and forwards them based on BIFT and BitString. Encapsulation of Forwarding Label is not in the scope of this document.

If there is no member in the multicast group, the PCE should send a PCUpd message with FI object to the ingress node, where F flag is not set.

8.4. Receiver information synchronization

Upon receiving multicast packets from a particular multicast group, egress node will synchronize the number of receivers in this multicast group with the PCE via PCRpt message with MRS object periodically.

After sending a PCUpd message with FI object to an ingress node for a particular multicast group, the PCE will synchronize the total number of receivers in this multicast group with the ingress node via PCUpd message with MRS object periodically.

If there is no member in the multicast group, the synchronization of receiver number information ends.

9. Deployment Considerations

10. Security Considerations

11. IANA Considerations

11.1. BIER-MULTICAST-CAPABILITY

IANA is requested to allocate a new code point within registry "STATEFUL-PCE-CAPABILITY TLV Flag Field" under "Path Computation Element Protocol (PCEP) Numbers" as follows:

Value	Description	Reference
TBD1	BIER-MULTICAST-CAPABILITY	This document

11.2. PCEP-ERROR Object

IANA is requested to allocate code-points in the "PCEP-ERROR Object Error Types and Values" subregistry for the following new error-type and error-value:

Error-Type	Description	Reference
10	Error-value = TBD2 B bit is not set	This document

11.3. New Objects

IANA is requested to allocate the following Object-Class Values in the "PCEP Objects" subregistry under the "Path Computation Element Protocol (PCEP) Numbers" registry:

Object-Class Value	Description	Reference
TBD3	Multicast Receiver Information	This document
TBD7	Multicast Receiver Information	This document
TBD9	Forwarding Indication	This document
TBD10	Multicast Receiver Status	This document

11.4. New TLVs

IANA is requested to allocate the following Object-Class Values in the "PCEP Objects" subregistry under the "Path Computation Element Protocol (PCEP) Numbers" registry:

Type	Description	Reference
TBD4	Multicast Source Address	This document
TBD5	Multicast Group Address	This document
TBD6	BIER Information TLV	This document
TBD8	VPN Information	This document

12. Contributor

13. Acknowledgement

14. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2362] Estrin, D., Farinacci, D., Helmy, A., Thaler, D., Deering, S., Handley, M., Jacobson, V., Liu, C., Sharma, P., and L. Wei, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification", RFC 2362, DOI 10.17487/RFC2362, June 1998, <https://www.rfc-editor.org/info/rfc2362>.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

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

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/info/rfc8231>.

[RFC8232] Crabbe, E., Minei, I., Medved, J., Varga, R., Zhang, X., and D. Dhody, "Optimizations of Label Switched Path State Synchronization Procedures for a Stateful PCE", RFC 8232, DOI 10.17487/RFC8232, September 2017, <https://www.rfc-editor.org/info/rfc8232>.

[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>.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <https://www.rfc-editor.org/info/rfc8281>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., 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 2018, <https://www.rfc-editor.org/info/rfc8296>.

Authors' Addresses

Huanan Li
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: lihn6@foxmail.com

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: wangaj3@chinatelecom.cn

Huaimo Chen
Futurewei
Boston
USA

Email: Huaimo.chen@futurewei.com

Ran Chen
ZTE Corporation
50 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: chen.ran@zte.com.cn