Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: March 26, 2015

IJ. Wijnands, Ed.
E. Rosen, Ed.
Cisco Systems, Inc.
A. Dolganow
Alcatel-Lucent
J. Tantsura
Ericsson

September 22, 2014

Encapsulation for Bit Index Explicit Replication in MPLS Networks
draft-wijnands-mpls-bier-encapsulation-00

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. When a multicast data packet enters the domain, the ingress router determines the set of egress routers to which the packet needs to be sent. The ingress router then encapsulates the packet in a BIER header. The BIER header contains a bitstring in which each bit represents exactly one egress router in the domain; to forward the packet to a given set of egress routers, the bits corresponding to those routers are set in the BIER header. The details of the encapsulation depend on the type of network used to realize the multicast domain. This document specifies the BIER encapsulation to be used in an MPLS network.

Table of Contents

1. Introduction

[BIER_ARCH] describes a new architecture for the forwarding of multicast data packets. That architecture provides optimal forwarding of multicast data packets through a "multicast domain". However, it does not require any explicit tree-building protocol, and does not require intermediate nodes to maintain any per-flow state. That architecture is known as "Bit Index Explicit Replication" (BIER).

This document will use terminology defined in [BIER_ARCH].

A router that supports BIER is known as a "Bit-Forwarding Router" (BFR). A "BIER domain" is a connected set of Bit-Forwarding Routers (BFRs), each of which has been assigned a BFR-prefix. A BFR-prefix is a routable IP address of a BFR, and is used by BIER to identify a BFR. A packet enters a BIER domain at an ingress BFR (BFIR), and leaves the BIER domain at one or more egress BFRs (BFERs). Each BFER must have a BFR-id as well as a BFR-prefix. A BFR-id is just a

number in the range [1,65535] that, within a BIER domain, identifies a BFER uniquely.

As described in [BIER_ARCH], BIER requires that multicast data packets be encapsulated with a header that carries the information needed to support the BIER forwarding procedures. This information includes a Set-Id (SI) and a BitString. Together the SI and the BitString identify the set of BFERs to which the packet must be delivered.

This document is applicable when a given BIER domain is both an IGP domain and an MPLS network. In this environment, the BIER encapsulation consists of two components:

- an MPLS label (which we will call the "BIER-MPLS label"); this label appears at the bottom of a packet's MPLS label stack.

- a BIER header, as specified in Section 3.

Following the BIER header is the "payload". The payload may be an IPv4 packet, an IPv6 packet, an ethernet frame, or an MPLS packet. If it is an MPLS packet, then an MPLS label stack immediately follows the BIER header. The top label of this MPLS label stack may be either a downstream-assigned label ([RFC3032]) or an upstream-assigned label ([RFC5331]. The BIER header contains information identifying the type of the payload.

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 RFC 2119 [RFC2119].

2. The BIER-MPLS Label

As stated in [BIER_ARCH], when a BIER domain is also an IGP domain, IGP extensions can be used by each BFR to advertise the BFR-id and BFR-prefix. The extensions for OSPF are given in [BIER_OSPF].

When a particular BIER domain is both an IGP domain and an MPLS network, we assume that each BFR will also use IGP extensions to advertise a set of one or more "BIER-MPLS" labels. When the domain contains a single "routing underlay" (see [BIER_ARCH]), a given BFR needs to advertise one such label for each SI. If the domain contains multiple routing underlays, a given BFR needs to advertise one such label per SI per each underlay in which that BFR has adjacencies.

The BIER-MPLS labels are locally significant (i.e., unique only to the BFR that advertises them) downstream-assigned MPLS labels. For

example, suppose there is a single routing underlay, the network is using a BitStringLength of 255, and that all BFERs in the domain have BFR-ids in the range [1,512]. Since each BIER BitString is 256 bits long, this requires the use of two SIs: SI=0 and SI=1. So each BFR will advertise, via IGP extensions, two MPLS labels for BIER: one corresponding to SI=0 and one corresponding to SI=1.

When a BFR receives an MPLS packet with one of its BIER-MPLS labels at the top of the stack, it will assume that a BIER header (see Section 3) immediately follows the stack. It will also infer the packet's SI from the label.

3. BIER Header

The BIER header is shown in Figure 1. This header appears after the end of the MPLS label stack, immediately after the MPLS-BIER label.

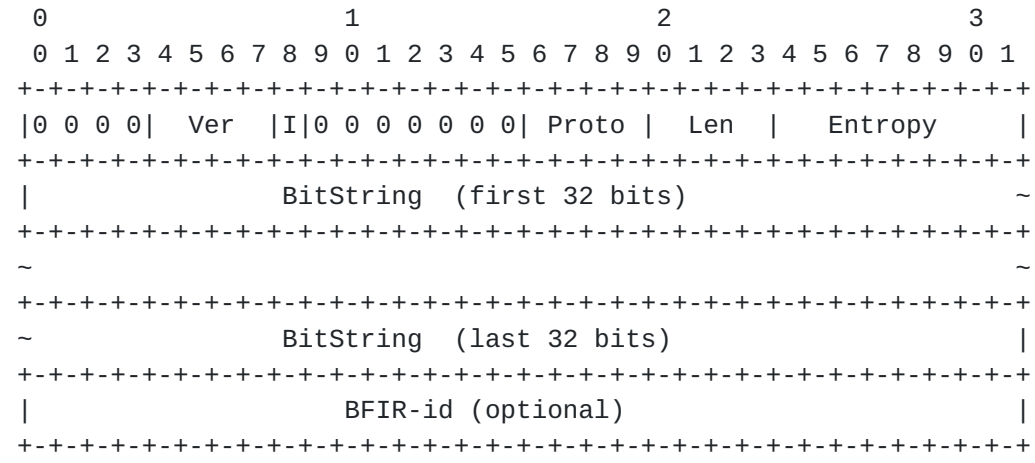

Figure 1: BIER Header

First nibble:

The first 4 bits of the header are all set to zero; this ensures that the BIER header will not be confused with an IP header.

Ver:

This 4-bit field identifies the version of the BIER header. This document specifies version 0 of the BIER header.

I:

This bit is set if and only if header contains the BFIR-id field. If this bit is clear, the header does not contain the BFIR-id, and the header ends at the end of the BitString.

Proto:

This 4-bit field identifies the type of the payload. (The "payload" is the packet or frame immediately following the BIER header.) The protocol field may take any of the following values:

1: MPLS packet with downstream-assigned label at top of stack.

2: MPLS packet with upstream-assigned label at top of stack (see [RFC5331]). If this value of the Proto field is used, the I bit MUST be set, and the BFR-id of the BFIR must be placed in the BFIR-id field. The BFIR-id provides the "context" in which the upstream-assigned label is interpreted.

3: Ethernet frame.

4: IPv4 packet.

6: IPv6 packet.

Len:

This 4-bit field encodes the length in bits of the BitString. If k is the length of the BitString, the value of this field is log2(k)-5. However, only certain values are supported:

1: 64 bits

2: 128 bits

3: 256 bits

4: 512 bits

5: 1024 bits

6: 2048 bits

7: 4096 bits

All other values of this field are illegal.

Entropy:

This 8-bit field specifies an "entropy" value that can be used for load balancing purposes. The BIER forwarding process may do equal cost load balancing, but the load balancing procedure MUST choose the same path for any two packets have the same entropy value.

If a BFIR is encapsulating (as the payload) MPLS packets that have entropy labels, the BFIR MUST ensure that if two such packets have the same MPLS entropy label, they also have the same value of the BIER entropy field.

BitString:

The BitString that, together with the packet's SI, identifies the destination BFERs for this packet. Note that the SI for the packet is inferred from the BIER-MPLS label that precedes the BIER header.

BFIR-id

This is the BFR-id of the BFIR. The BFR-id is encoded in the 32-bit field as an unsigned integer in the range [1,65535].

This field is optional, and is present only when the I bit is set.

4. Imposing and Processing the BIER Encapsulation

When a BFIR receives a multicast packet from outside the BIER domain, it carries out the following procedure:

1. By consulting the "multicast flow layer" ([BIER_ARCH]), it determines the value of the "Proto" field.

2. By consulting the "multicast flow layer", it determines the set of BFERs that must receive the packet.

3. The BFIR looks up the BFR-id of each of those BFERs.

4. The BFIR converts each such BFR-id into (SI, BitString) format, as described in [BIER_ARCH].

5. All such BFR-ids that have the same SI can be encoded into the same BitString. Details of this encoding can be found in [BIER_ARCH]. For each distinct SI that occurs in the list of the packet's destination BFERs:

   a. The BFIR make a copy of the multicast data packet, and encapsulates the copy in a BIER header (see Section 3). The BIER header contains the BitString that represents all the destination BFERs whose BFR-ids correspond to the given SI.

   b. If the payload is an MPLS packet whose label stack begins with an upstream-assigned label, the BFIR-id field MUST be present. (Whether a particular MPLS packet payload begins

with an upstream-assigned label is learned from the multicast flow layer.) The BFIR-id MAY be included in other cases as well.

c. The BFIR then applies to that copy the forwarding procedure of [BIER_ARCH]. This may result in one or more copies of the packet (possibly with a modified BitString) being transmitted to a neighboring BFR.

d. Before transmitting a copy of the packet to a neighboring BFR, the BFIR finds the BIER-MPLS label that was advertised by the neighbor as corresponding to the given SI. An MPLS label stack is then preprended to the packet. This label stack [RFC3032] will contain one label, the aforementioned BIER-MPLS label. The "S" bit MUST be set, indicating the end of the MPLS label stack. The packet may then be transmitted to the neighboring BFR. (This may result in additional MPLS labels being pushed on the stack. For example, if an RSVP-TE tunnel is used to transmit packets to the neighbor, a label representing that tunnel would be pushed onto the stack.)

When an intermediate BFR receives a packet with one of its own BIER-MPLS labels at the top of the label stack, it infers the SI from the label, and follows the forwarding procedures of [BIER_ARCH]. If it forwards a copy of the packet to a neighboring BFR, it first swaps the label at the top of the label stack with the BIER-MPLS label that the neighbor advertised that corresponds to the same SI and routing underlay.

Thus a BIER-encapsulated packet in an MPLS network consists of a packet that has:

- An MPLS label stack with a BIER-MPLS label at the bottom of the stack.
- A BIER header, as described in Section 3.
- The payload, a multicast data packet.

The payload may be an IPv4 packet, an IPv6 packet, an ethernet frame, or an MPLS packet. If it is an MPLS packet, the BIER header is followed by a second MPLS label stack; this stack is separate from the stack that precedes the BIER header.

5. IANA Considerations

This document has no actions for IANA.

6. Security Considerations

As this document makes use of MPLS, it inherits any security considerations that apply to the use of the MPLS data plane.

As this document makes use of IGP extensions, it inherits any security considerations that apply to the IGP.

The security considerations of [BIER ARCH] also apply.

7. Acknowledgements

The authors wish to thank Rajiv Asati, John Bettink, Nagendra Kumar, Christian Martin, Neale Ranns, Greg Shepherd, and Ramji Vaithianathan for their ideas and contributions to this work.

8. Contributor Addresses

Below is a list of other contributing authors in alphabetical order:

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
Antwerp 2018
Belgium

Email: wim.henderickx@alcatel-lucent.com

Martin Horneffer
Deutsche Telekom
Hammer Str. 216-226
Muenster 48153
DE

Email: Martin.Horneffer@telekom.de

Uwe Joorde
Deutsche Telekom
Hammer Str. 216-226
Muenster D-48153
DE

Email: Uwe.Joorde@telekom.de

Tony Przygienda
Ericsson

Email: antoni.przygienda@ericsson.com

9. Normative References

[BIER_ARCH] Wijnands, IJ., "Multicast using Bit Index Explicit Replication Architecture", internet-draft draft-wijnands-bier-architecture-00, September 2014.

[BIER_OSPF] Kumar, N., "OSPF Extension for Bit Index Explicit Replication", internet-draft draft-kumar-ospf-bier-extension-00, September 2014.

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

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, August 2008.

Authors' Addresses

IJsbrand Wijnands (editor)
Cisco Systems, Inc.
De Kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Eric C. Rosen (editor)
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, Massachusetts 01718
USA

Email: erosen@cisco.com

Andrew Dolganow
Alcatel-Lucent
600 March Rd.
Ottawa, Ontario K2K 2E6
Canada

Email: andrew.dolganow@alcatel-lucent.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, California 95134
USA

Email: jeff.tantsura@ericsson.com