

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 31, 2014

J. Guichard
C. Pignataro, Ed.
S. Spraggs
S. Bryant
Cisco
September 27, 2013

Carrying Metadata in MPLS Networks
draft-guichard-sfc-mpls-metadata-00

Abstract

This document defines the mechanism for identifying the presence of metadata carried in addition to the payload in MPLS packets.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://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 March 31, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

Internet-Draft

MPLS Metadata

September 2013

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
1.1.	Terminology	3
2.	Metadata Component Structure	3
3.	Metadata Channel Header Format	4
3.1.	Metadata Encapsulation Format	4
4.	Load-balancing Considerations	5
5.	Metadata and MPLS Label Stack	5
6.	Data Plane Processing of MPLS Packets Containing Metadata . .	6
6.1.	Egress LSR	6
6.2.	Ingress LER/LSR	6
6.3.	Transit LSR	7
6.4.	Penultimate Hop LSR	7
7.	Data Plane Processing of MPLS Packets Containing Metadata . .	7
8.	IANA Considerations	8
9.	Security Considerations	8
10.	Contributing Authors	8
11.	Acknowledgments	8
12.	References	9
12.1.	Normative References	9
12.2.	Informative References	9
Appendix A.	Alternative Options	10
	Authors' Addresses	10

[1.](#) Introduction

This document defines the mechanism for identifying the presence of metadata carried in addition to the payload in MPLS packets. The metadata header is common across all encapsulations (including IPv4, IPv6, and MPLS) and is defined in [[I-D.guichard-metadata-header](#)].

[1.1.](#) Terminology

ACH Associated Channel Header

AL Application Label

EL Entropy Label

ELI Entropy Label Indicator

G-ACH Generic Associated Channel

GAL Generic Associated Channel Label

TL Top Label

MCH Metadata Channel Header

MD Metadata

[2.](#) Metadata Component Structure

The addition of metadata to packets enables the instrumentation of user packets, and service chaining, although it is anticipated that the ability to allow packets to carry metadata of use to the infrastructure and specific handling instructions will enable other uses.

Metadata carried within an MPLS packet is prefaced by a Metadata Channel Header (MCH) as defined in [[I-D.guichard-metadata-header](#)], with the first nibble of the MCH set to 0000b.

Metadata is distinguished from IP payloads using similar methods to those developed in pseudowires and MPLS-TP [[RFC4385](#)] [[RFC5586](#)].

Two scenarios are presented for MPLS environments where metadata may be required.

1. IPv4, IPv6 or pseudo-wire payload. In this case the metadata will be carried within MPLS packets between the MCH and the

original MPLS payload. A GAL reserved label [[RFC5586](#)] is used to indicate that metadata is carried within the MPLS packet and that an MCH immediately follows the bottom of the label stack.

2. MPLS payload. In this case a new label stack will be created for the section over which the metadata is relevant and the original MPLS packet (MPLS label stack and MPLS payload) will be carried in the payload section described below. An example where this type of scenario may be required is when a hierarchical LSP needs to be instrumented. In this case, rather than pushing the labels associated with the hierarchical section onto the existing label stack, the original label stack is preserved and placed along with its associated payload in the payload section described below. A new label stack, indicating the presence of metadata (by way of the GAL), the MCH, and the metadata itself is then built for the MPLS section requiring instrumentation and sent.

[3.](#) Metadata Channel Header Format

The presence of metadata within an MPLS packet must be indicated in the encapsulation. This document defines that the G-ACh Generic Associated Channel Label (GAL) [[RFC5586](#)] with label value 13 is utilized for this purpose. The GAL label provides a method to identify that a packet contains an "Associated Channel Header (ACH)" followed by a non-service payload.

[[RFC5586](#)] identifies the G-ACh Generic Associated Channel by setting the first nibble of the ACH that immediately follows the bottom label

in the stack if the GAL label is present, to 0001b. Further [RFC5586] expects that the ACH not be used to carry user data traffic. This document proposes an extension to allow the first nibble of the ACH to be set to 0000b and, when following the GAL, be interpreted using the semantics defined in [I-D.guichard-metadata-header] to allow metadata to be carried through the G-ACH channel.

The metadata is distinguished from OAM by the use of 0000b in the first nibble. This is consistent with the practise developed in pseudowire [RFC4928] which uses a first nibble of 0000b to indicate the presence of information to be used by the forwarding plane to correctly forward the packet (i.e. the PW control word [RFC4385]).

3.1. Metadata Encapsulation Format

Figure 1 depicts the Metadata encapsulation format:

Guichard, et al. Expires March 31, 2014 [Page 4]

Internet-Draft MPLS Metadata September 2013

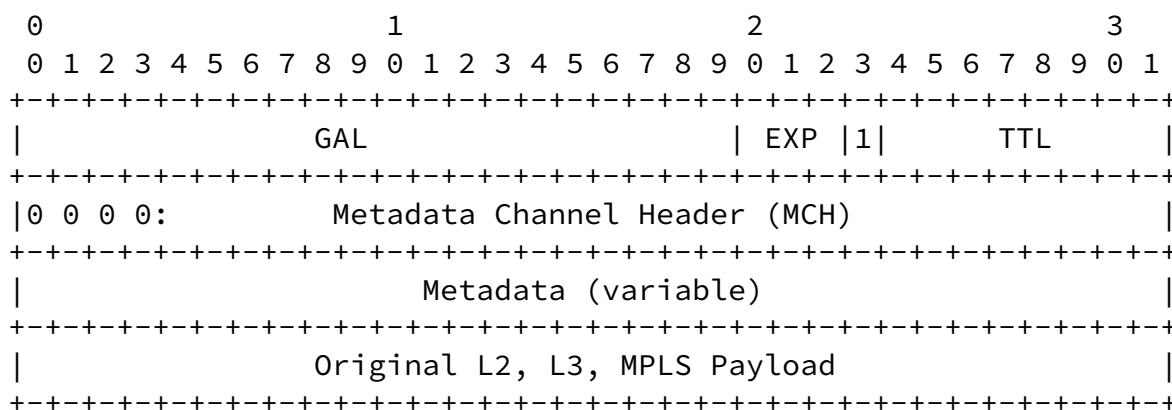


Figure 1: Metadata Encapsulation Format

The meanings of the fields in the metadata packet format are as follows:

- o The GAL (reserved label of value 13) is used to indicate that an ACH or MCH appears immediately after the bottom of the label stack. The first nibble of the channel header is used to identify whether the format is to be interpreted as an ACH or MCH.

- o If the first nibble is set as 0000b, this indicates that an MCH will sit beneath the label stack.
- o Immediately following the MCH will be the metadata. The length and format of the metadata is outside the scope of this document and will vary depending upon the "Metadata Channel Type" specified in the MCH.
- o Beneath the metadata will be the original packet payload. This could be L2, L3 or MPLS payload.

4. Load-balancing Considerations

The approach in this document is consistent with the use of utilizing 0000b as the first nibble after the MPLS label stack, as described in [[RFC4928](#)]. In this case, the MCH starts with 0000b. Load balancing is achieved utilizing the entropy label and following the methods defined in [[RFC6790](#)].

5. Metadata and MPLS Label Stack

Only one piece of metadata can be carried for each payload (L2, L3, or MPLS). As a consequence there MUST be only one GAL label in the label stack. Entropy labels MAY be present in the label stack but

they MUST be indicated using the Entropy Label Indicator (ELI) as described in [[RFC6790](#)].

6. Data Plane Processing of MPLS Packets Containing Metadata

6.1. Egress LSR

Suppose egress LSR Y is capable of processing metadata. LSR Y indicates this to all other LER's and LSR's via signaling (see [Section 7](#) for more discussion on this subject) or through direct configuration.

LSR Y MUST be prepared to process packets carrying metadata and those without. If a GAL is present in the MPLS label stack, the receiving

LSR MUST inspect the first nibble after the end of the label stack to identify the presence of an MCH or an ACH, and process the packet accordingly. An LSR SHOULD NOT push a GAL on a packet that does not contain an MCH or an ACH.

If a particular LER or LSR chooses to send traffic without metadata, LSR Y's processing of the received label stack (which might be empty) and payload is based on normal MPLS processing rules. If LER/LSR X chooses to send metadata, then LSR Y will receive an MPLS packet constructed as follows:

<Top Label (TL), AL, GAL> <MCH> <metadata> <remaining packet payload>

LSR Y recognizes TL as the label it distributed to its upstream LER/LSR and pops the TL (note that the TL signalled by LSR Y may be an implicit null label, in which case it doesn't appear in the label stack and LSR Y MUST process the packet starting with the AL label (if present) and/or the GAL label.) LSR Y recognizes the GAL with S bit set. LSR Y then processes the metadata, starting with the MCH (0000b), which will determine how LSR Y processes the underlying payload.

[6.2.](#) Ingress LER/LSR

If an egress LSR Y indicates via signaling or through direct configuration of other LER's/LSR's that it can process metadata, the steps that Ingress LER/LSR X performs to insert metadata are as follows:

1. On an incoming packet, identify the application to which the packet belongs and from this the egress LSR; based on these two components determine whether metadata needs to be added to the

incoming packet.

2. For packets requiring the insertion of metadata, build the appropriate MCH and prepend the metadata and the MCH to the existing payload; then, push the GAL label on to the label stack with the S bit set. For packets not requiring insertion of metadata, this step is a NOOP.

3. Push the application label (AL) label (if required) on to the label stack.
4. If Entropy is required then pick appropriate fields as input to the load-balancing function; apply the load-balancing function to these input fields and generate the Entropy label (EL) value.
5. Push the EL and the ELI labels on to the label stack.
6. Determine the top label (TPL) and push it on top of the ELI and EL (if present). The ordering of the AL and the ELI plus EL pair is not critical other than that the egress LSR processing the ELI MUST process all remaining labels in the stack and the metadata. The S bit for the ELI and EL MUST be zero (i.e., S bit is not set). The TTL for the EL MUST be zero to ensure that it is not used inadvertently for forwarding. The TC for the EL may be any value.
7. Determine the output interface, and transmit.

[6.3.](#) Transit LSR

Transit LSRs may operate with no change in forwarding behavior.

[6.4.](#) Penultimate Hop LSR

No change is needed at penultimate hop LSRs.

[7.](#) Data Plane Processing of MPLS Packets Containing Metadata

Two levels of set-up are required to support metadata. The first is an indication that the device or LSP is capable of supporting metadata. This could be done either using the NMS or by using capabilities exchange mechanisms. For example an IGP ([RFC4971](#) in ISIS) or MPLS protocols such as [RFC5036](#), [RFC3209](#), or [RFC3107](#). The specific mechanism for signaling the support of metadata is outside the scope of this document and will be defined elsewhere.

The second set-up required is by the actual application using the

information contained in the metadata. Again this could be done

using either the NMS or a signaling protocol. It is anticipated this type of signaling is specifically associated with the application and would be specified elsewhere.

8. IANA Considerations

This document makes no request of IANA.

[Note to RFC Editor: this section may be removed on publication as an RFC.]

9. Security Considerations

The addition of metadata to a packet increases the amount of processing required by the LSR receiving the packet, and thus may be used in a denial of service attack vector. However MPLS networks carefully manage their adjacencies and only accept labeled packets from trusted neighbors. Provided this current level of neighbor verification remains in place no additional risk results.

The metadata itself may be an attack vector with either the originating LSR or a man in the middle inserting malicious content. The trust model of the MPLS network itself, described earlier in this section guards against a man in the middle attack and ensures that the originating LER/LSR is a trusted party.

If the ingress LER/LSR is taking instructions from a third party in the specific metadata to insert, there MUST be a sufficient trust relationship between the ingress LER/LSR and the third party.

The security considerations associated with each metadata type MUST be specified as part of its definition.

10. Contributing Authors

- o Clarence Filsfils <cfilsfil@cisco.com>
- o Dan Frost <danfrost@cisco.com>

11. Acknowledgments

The authors would like to thank Giles Heron and Tom Nadeau for their review and useful comments.

[12.](#) References

[12.1.](#) Normative References

[I-D.guichard-metadata-header]

Guichard, J., Spraggs, S., Pignataro, C., and S. Bryant, "Common Metadata Header Format for IP/MPLS Networks", [draft-guichard-metadata-header-00](#) (work in progress), June 2013.

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

[12.2.](#) Informative References

[I-D.kompella-mpls-special-purpose-labels]

Kompella, K. and A. Farrel, "Allocating and Retiring Special Purpose MPLS Labels", [draft-kompella-mpls-special-purpose-labels-04](#) (work in progress), May 2013.

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", [RFC 3107](#), May 2001.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", [RFC 4385](#), February 2006.

[RFC4928] Swallow, G., Bryant, S., and L. Andersson, "Avoiding Equal Cost Multipath Treatment in MPLS Networks", [BCP 128](#), [RFC 4928](#), June 2007.

[RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", [RFC 4971](#), July 2007.

[RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", [RFC 5036](#), October 2007.

[RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", [RFC 5586](#), June 2009.

[RFC 6790](#), November 2012.

[Appendix A](#). Alternative Options

This appendix lists alternative options for metadata indication that were considered but ultimately discarded:

- o Starting the MCH with the first nibble as 0010b. This first nibble is overloading the IP version field, and thus the creation of new first nibbles needs to be a conservative process, since each new nibble used for other purposes prevents that nibble being used to identify a new IP type at some time in the future.
- o Extending a G-ACh to be able to carry user data. This has been discussed at length within the IETF and it seems the consensus is this structure should not carry customer payload.
- o Assign a new reserved label, either directly or as an extension label as proposed in [[I-D.kompella-mpls-special-purpose-labels](#)] to indicate the presence of metadata. In the first case it utilizes another reserved label, which are becoming sparse. In the second case it increases the size of the label stack.

The method described in this document has more benefits and fewer drawbacks than these three.

Authors' Addresses

Jim Guichard
Cisco Systems, Inc.

Email: jguichar@cisco.com

Carlos Pignataro (editor)
Cisco Systems, Inc.
7200-12 Kit Creek Road

Research Triangle Park, NC 27709
US

Email: cpignata@cisco.com

Guichard, et al.

Expires March 31, 2014

[Page 10]

Internet-Draft

MPLS Metadata

September 2013

Simon Spraggs
Cisco Systems, Inc.
10 New Square Park
Bedfont Lakes, Feltham TW14 8HA
United Kingdom

Email: sspraggs@cisco.com

Stewart Bryant
Cisco Systems, Inc.
10 New Square Park
Bedfont Lakes, Feltham TW14 8HA
United Kingdom

Email: stbryant@cisco.com

Guichard, et al.

Expires March 31, 2014

[Page 11]