Network Working Group
Internet-Draft
Updates: RFC 3032, RFC 3270, RFC 5129, RFC 3272, RFC 3443, RFC 3469, RFC 3564, RFC 3985, RFC 4182, RFC 4364, RFC 4379, RFC 4448, RFC 4761 (if approved)
Intended status: Standards Track
Expires: June 8, 2009

L. Andersson
Acreo AB
R. Asati
Cisco Systems

December 5, 2008

Multi-Protocol Label Switching (MPLS) label stack entry: "EXP" field renamed to "Traffic Class" field
draft-ietf-mpls-cosfield-def-08.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on June 8, 2009.

Abstract

The early Multiprotocol Label Switching (MPLS) documents defined the form of the MPLS Label Stack entry. This includes a three bit field called the "EXP field". The exact use of this field was not defined by these documents, except to state that it was to be "reserved for experimental use".

Although the intended use of the EXP field was as a "Class of Service" (CoS) field, it was not named the CoS field by these early documents because the use of such a CoS field was not considered to be sufficiently defined. Today a number of standards documents define its usage as a CoS field. .

To avoid misunderstanding about how this field may be used, it has become increasingly necessary to rename this field. This document changes the name of the field to the "Traffic Class field" ("TC field".) In doing so it also updates documents that define the current use of the EXP this field.

Table of Contents

1. Introduction

The format of a MPLS label stack entry is defined by RFC 3032 [RFC3032], include a three bit field called the "EXP field". The exact use of this field is not defined by RFC 3032 leaves, except to state that it is to be "reserved for experimental use".

The EXP field, from the start, was intended to carry "Class of Service" (CoS) information. The field was actually called the "Class of Service field" in the early versions of the working group document that was published as RFC 3032. However at the time that RFC 3032 was published the exact usage of this "Class of Service" field was not agreed and the field was designated as "Experimental use"; hence the name has since then been the "EXP Field".

The designation "for Experimental use" has led other Standards Development Organizations (SDO) and implementors to the assume that it possible to use the field for other purposes. This document changes the name of the field to clearly indicate its use as a traffic classification field.

At first we discussed to use the original "CoS field" as the name for the field, but it has been pointed that this name does not cover the following changes with respect to its usage, since RFC 3032 was published.

1. The use of the EXP field was first defined in RFC 3270 [RFC3270] where a method to define a variant of DiffServ Label Switched Paths (LSP) called EXP-Inferred-PSC LSP (E-LSPs) was specified.

The PSC is a two stage acroynym that is expanded as Per Hop Behavior (PHB) and PHB Scheduling Class (PSC).

2. The use of the EXP field as defined in RFC 3270 has been further extended in RFC 5129 [RFC5129], where methods for explicit congestion marking in MPLS are defined.

This document, hence, uses the name "Traffic Class Field (TC Field)", which better covers the potential use. The MPLS TC field relates to an MPLS encapsulated packet the same way as the IPv6 TC field relates to an IPv6 encapsulted packet or the IPv4 Precedence field relates to an IPv4 encapsulated packet.

The definitions of how the EXP field is used are perfectly clear in RFC 3270 and RFC 5129. However, these RFCs do not explicitly state they update RFC 3032, and this fact was not captured in the RFC repository until after the work on this document were started. This document updates RFC 3032, RFC 3270 and RFC 5129 to clarify the intended usage of the TC field. Section 2 explains the changes.

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. Details of change

The three RFCs are now updated according to the following.

2.1. RFC 3032

RFC 3032 states on page 4:

> 3. Experimental Use
>
> This three-bit field is reserved for experimental use.

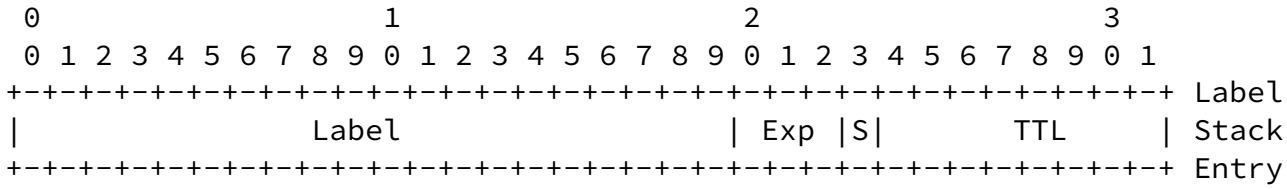
This paragraph is now changed to:

> 3. Traffic Class (TC) field

This three-bit field is used to carry Traffic Class information and the change of the name is applicable to all places it occurs in IETF RFCs and other IETF documents.

RFC 3270 and RFC 5129 updates the definition of the TC field and describes how to use the field.

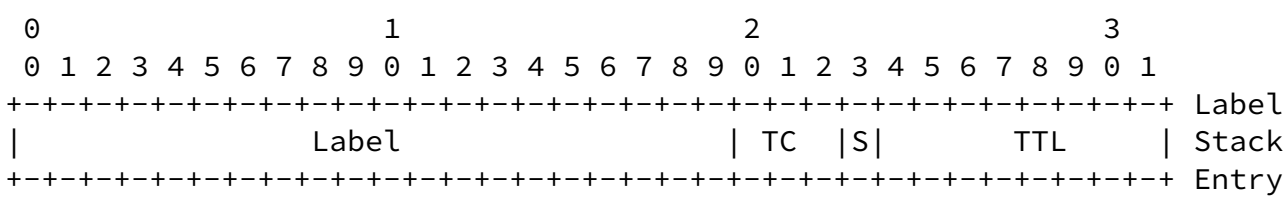
In Figure 1 on page 3 in RFC3032 the format of a label stack entry is specified as:



Label: Label Value, 20 bits
Exp: Experimental Use, 3 bits
S: Bottom of Stack, 1 bit
TTL: Time to Live, 8 bits

Figure 1

Figure 1 in RFC 3032 is now changed to match the change of name of the TC field to:



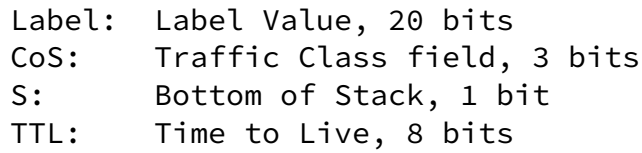


Figure 1 (new)

Note: The designation of the picture above as "Figure 1 new" is introduced as a way to distinguish the figures in this draft. It will still be "Figure 1." in RFC 3032.

2.2. RFC 3270

RFC 3270 says on page 6:

1.2 EXP-Inferred-PSC LSPs (E-LSP)

> A single LSP can be used to support one or more OAs. Such LSPs can support up to eight BAs of a given FEC, regardless of how many OAs these BAs span. With such LSPs, the EXP field of the MPLS Shim Header is used by the LSR to determine the PHB to be applied to the packet. This includes both the PSC and the drop preference.
>
> We refer to such LSPs as "EXP-inferred-PSC LSPs" (E-LSP), since the PSC of a packet transported on this LSP depends on the EXP field value for that packet.
>
> The mapping from the EXP field to the PHB (i.e., to PSC and drop precedence) for a given such LSP, is either explicitly signaled at label set-up or relies on a pre-configured mapping.
>
> Detailed operations of E-LSPs are specified in section 3 below.

RFC 3270 is now updated like this:

a. A new paragraph is added at the end of section 1 "Introduction":

The EXP field has been renamed to the TC field, and thus all references in RFC 3270 to EXP field SHOULD be taken to refer to the TC field.

b. A new term is added to section 1.1 "Terminology":

TC Traffic Class (replaces the term EXP)

c. In section 1.1 "Terminology" the acronym E-LSP is now understood to mean :

E-LSP Explicitly TC-encoded-PSC LSP

Section 1.2 on page 5 in RFC 3270 is now changed to:

1.2 Explicitly TC-encoded-PSC LSPs (E-LSP)

The EXP field has been renamed to the TC field, and thus all references in RFC 3270 to EXP field SHOULD be taken to refer to the TC field. However, we retain the acronym E-LSP (Explicitly TC-encoded-PSC LSP) as the acronym is in widespread use.

A single LSP can be used to support one or more OAs. Such LSPs can support up to eight BAs of a given FEC, regardless of how many OAs these BAs span. With such LSPs, the TC field of the MPLS Shim Header is used by the LSR to determine the PHB to be applied to the packet. This includes both the PSC and the drop preference.

We refer to such LSPs as "Explicitly TC-encoded-PSC LSP" (E-LSP), since the PSC of a packet transported on this LSP depends on the TC field (previously called the EXP field) value for that packet.

The mapping from the TC field to the PHB (i.e., to PSC and drop precedence) for a given such LSP, is either explicitly signaled at label set-up or relies on a pre-configured mapping.

This is an update to RFC 3032 [RFC3032] in line with the original intent of how this field in the MPLS Shim Header should be used (as TC field). The RFC 3270 has itself been updated by RFC 5129 [RFC5129].

Detailed operations of E-LSPs are specified in section 3 of RFC3270.

2.3. RFC 5129

RFC 5129 is now updated like this:

A new paragraph is added at the end of section 1.1 "Background":

> The EXP field has been renamed to the TC field, and thus all references in RFC 5129 to EXP field SHOULD be taken to refer to the TC field.

Section 2 (bullet 3) on page 6 of RFC 5129 says:

> o A third possible approach was suggested by [Shayman]. In this scheme, interior LSRs assume that the endpoints are ECN-capable, but this assumption is checked when the final label is popped. If an interior LSR has marked ECN in the EXP field of the shim header, but the IP header says the endpoints are not ECN-capable, the edge router (or penultimate router, if using penultimate hop popping) drops the packet. We recommend this scheme, which we call `per-domain ECT checking', and define it more precisely in the following section. Its chief drawback is that it can cause packets to be forwarded after encountering congestion only to be dropped at the egress of the MPLS domain. The rationale for this decision is given in Section 8.1.

Section 2 (bullet 3) of RFC 5129 is now updated changed to:

> o A third possible approach was suggested by [Shayman]. In this scheme, interior LSRs assume that the endpoints are ECN-capable, but this assumption is checked when the final label is popped. If an interior LSR has marked ECN in the TC field of the shim header, but the IP header says the endpoints are not TC-capable, the edge router (or penultimate router, if using penultimate hop popping) drops the packet. We recommend this scheme, which we call `per-domain ECT checking', and define it more precisely in the following section. Its chief drawback is that it can cause packets to be forwarded after encountering congestion only to be dropped at the egress of the MPLS domain. The rationale for this decision is given in Section 8.1. This scheme is an update to RFC 3032 [RFC3032] and RFC 3270 [RFC3270].

2.4. The Scope of this Change

There are several places in the RFCs that has explicitly updated by this document that reference the "Exp field", sometimes they refer to the field as "Exp bits", "EXP bits" and "EXP". In all those

instances the references SHOULD be taken to reference the TC field.

There are also other RFCs, e.g. RFC 3272 [RFC3272], RFC 3443 [RFC3443], RFC 3469 [RFC3469], RFC 3564 [RFC3564], RFC 3985 [RFC3985], RFC 4182 [RFC4182], RFC 4364 [RFC4364], RFC 4379 [RFC4379], RFC 4448 [RFC4448] and RFC 4761 [RFC4761] that references the "Exp field", sometimes they refer to the field as "Exp bits", "EXP bits" and "EXP". For all RFCs, including but not limited to those mentioned in this paragraph, such references SHOULD be taken to reference the TC field.

3. Use of the TC field

Due to the limited number of bits in the TC field, their use for QoS and ECN functions is intended to be flexible. These funtions may rewrite all or some of the bits in the TC field.

Current implementations look at the TC field with and without label context and the TC field may be copied to the label stack entries that are pushed onto the label stack. This is done to avoid that label stack entries that are pushed on to an existing label stack have different TF fields from the rest of the label stack entries.

4. IANA considerations

There are no requests for IANA allocation of code points in this document.

5. Security considerations

This document only changes the name of one field in the MPLS Shim Header and thus does not introduce any new security considerations.

6. Acknowledgments

The author would like to thank Stewart Bryant, Bruce Davie, George Swallow, and Francois Le Faucheur for their input to and review of the current document.

The author also like to thanks George Swallow, Khatri Paresh and Phil Bedard for their help with grammar and spelling, and a special thanks to Adrian Farrel for a careful review and help trawling the RFC-sea for RFCs that references the EXP field.

7. References

7.1. Normative References

[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.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

[RFC3272] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., and X. Xiao, "Overview and Principles of Internet Traffic Engineering", RFC 3272, May 2002.

[RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks", RFC 3443, January 2003.

[RFC3469] Sharma, V. and F. Hellstrand, "Framework for Multi-Protocol Label Switching (MPLS)-based Recovery", RFC 3469, February 2003.

[RFC3564] Le Faucheur, F. and W. Lai, "Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering", RFC 3564, July 2003.

[RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

[RFC4182] Rosen, E., "Removing a Restriction on the use of MPLS Explicit NULL", RFC 4182, September 2005.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[RFC4448] Martini, L., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

[RFC5129] Davie, B., Briscoe, B., and J. Tay, "Explicit Congestion Marking in MPLS", RFC 5129, January 2008.

7.2. Informative references

[Shayman] Shayman, M. and R. Jaeger, University of Michigan, "Using ECN to Signal Congestion Within an MPLS Domain", Work in Progress, November 2000.", <http://www.watersprings.org/pub/id/draft-shayman-mpls-ecn-00.txt/>.

Authors' Addresses

Loa Andersson
Acreo AB

Email: loa@pi.nu

Rajiva Asati
Cisco Systems

Email: rajiva@cisco.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at

http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.