MPLS
Internet-Draft
Intended status: Standards Track
Expires: September 3, 2015

S. Bryant
G. Swallow
S. Sivabalan
Cisco Systems

March 2, 2015

RFC6374 Synonymous Flow Labels
draft-bryant-mpls-synonymous-flow-labels-00

Abstract

This document describes a method of providing flow identification information when making RFC6374 performance measurements. This allows RFC6374 measurements to be made on multi-point to point LSPs and allows the measurement of flows within an MPLS construct using RFC6374.

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 September 3, 2015.

Table of Contents

1. Introduction

[I-D.bryant-mpls-flow-ident] describes the requirement for introducing flow identities when using RFC6374 [RFC6374] packet loss measurements. In summary RFC6374 uses the RFC6374 packet as the packet accounting demarcation point. Unfortunately this gives rise to a number of problems that may lead to significant packet accounting errors:

1. Where a flow is subjected to Equal Cost Multi-Path (ECMP) treatment packets may arrive out of order with respect to the RFC6374 packet.

2. Where a flow is subjected to ECMP treatment packets may arrive at different hardware interfaces, thus requiring reception of an RFC6374 packet on one interface to trigger a packet accounting action on another interface which may not be co-located with it. This is a difficult technical problem to address with the required degree of accuracy.

3. Even where there is no ECMP (for example on RSVP-TE, MPLS-TP LSPs and PWs) local processing may be distributed over a number of cores, leading to synchronization problems.

4. Some forwarder implementations have a long pipeline between processing a packet and incrementing the associated counter.

An approach to mitigating these synchronization issue is described in [I-D.tempia-opsawg-p3m] in which packets are batched by the sender and each batch is marked in some way such that adjacent batches can be easily recognized by the receiver.

An additional problem arises where the LSP is a multi-point to point LSP, since MPLS does not include a source address in the packet. Network management operations require the measurement of packet loss between a source and destination. It is thus necessary to introduce some source specific information into the packet to identify packet batches from a specific source.

This document describes a method of accomplishing this by using a technique called synonymous flow labels Section 3 in which labels which mimic the behaviour of other labels provide the packet batch identifiers and enable the per batch packet accounting.

2. Requirements Language

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 [RFC2119].

3. Synonymous Flow Labels

A synonymous flow label (SFL) is defined to be a label that causes exactly the same forwarding behaviour at the egress Label Switching Router (LSR) as another label, except that it also causes an additional agreed action to take place on the packet. In this application the agreed action is the recording of the receipt of the packet by incrementing a packet counter. This is a natural action in many MPLS implementations, and where supported this permits the implementation of high quality packet loss measurement without any change to the packet forwarding system.

Consider an MPLS application such as a pseudowire (PW), and consider that it is desired to use the approach specified in this document to make a packet loss measurement. By some method outside the scope of this text, two labels, synonymous with the PW labels are obtained from the egress terminating provider edge (T-PE). By alternating between these SLs and using them in place of the PW label, the PW packets may be batched for counting without any impact on the PW forwarding behaviour (note that strictly only one SL is needed in

this application, but that optimization is a matter for the implementor).

Now consider an MPLS application that is multi-point to point such as a VPN. Here it is necessary to identify a packet batch from a specific source. This is achieved by making the SLs source specific, so that batches from one source are marked differently from batches form another source. Note that the sources all operate independently and asynchronously from each other, independently co-ordinating with the destination.

Finally we need to consider the case where there is no MPLS application label such as occurs when sending IP over an LSP. In this case introducing an SL that was synonymous with the LSP label would introduce network wide forwarding state. This would not be acceptable for scaling reasons. We therefore have no choice but to introduce an additional label. Where penultimate hop popping (PHP) is in use the semantics of this additional label can be similar to the LSP label. Where PHP is not in use he semantics are similar to an MPLS explicit NULL. In both cases with the additional semantics of the SL.

Note that to achieve the goals set out in Section 1 SLs need to be allocated from the platform label table.

4. User Service Traffic in the Data Plane

As noted in Section 3 it is necessary to consider two cases:

1. Applications label present

2. Single label stack

4.1. Applications Label Present

Figure 1 shows the case in which both an LSP label and an application label is present in the MPLS label stack. Uninstrumented traffic runs over the "normal" stack, and instrumented flows run over the SFL stack with the SFL used to indicate the packet batch.

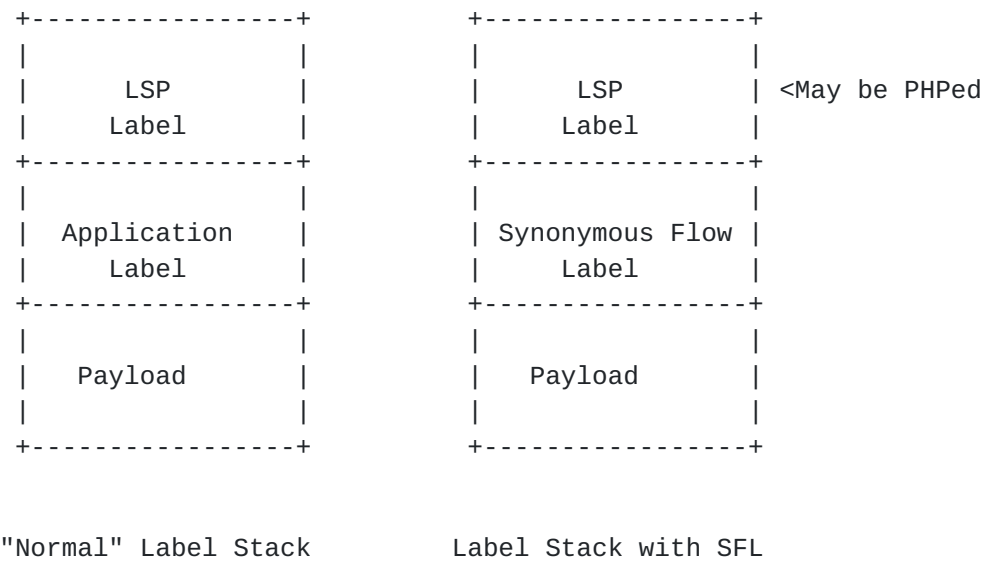

Figure 1: Use of Synonymous Labels In A Two Label MPLS Label Stack

At the egress LSR the LSP label is popped (if present). Then the SFL is processed in exactly the same way as the corresponding application label would have been processed. Where the SFL is being used to support RFC6374 packet loss measurements, as an additional operation, the total number of packets received with this particular SFL is recorded.

4.2. Single Label Stack

Figure 2 shows the case in which only an LSP label is present in the MPLS label stack. Uninstrumented traffic runs over the "normal" stack and instrumented flows run over the SFL stack with the SFL used to indicate the packet batch. However in this case it is necessary for the ingress LSR to first push the SFL and then to push the LSP label.

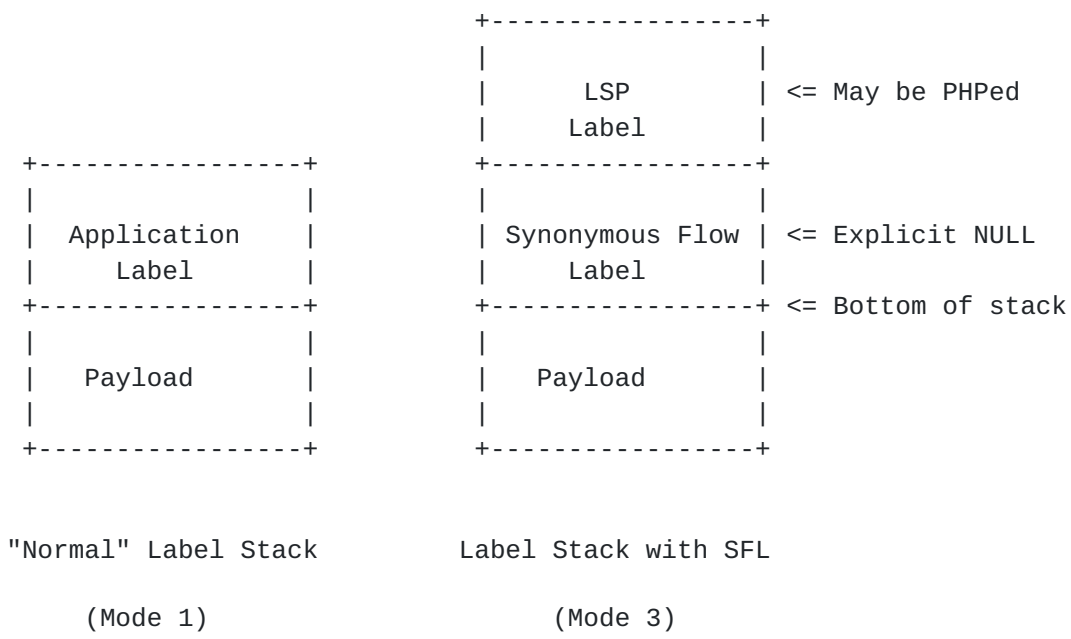

Figure 2: Use of Synonymous Labels In A Single Label MPLS Label Stack

At the receiving LSR it is necessary to consider two cases:

1. Where the LSP label is still present

2. Where the LSP label is penultimate hop popped

If the LSP label is present, it processed exactly as it would normally processed and then it is popped. This reveals the SFL which in the case of RFC6374 measurements is simply counted and then discarded. In this respect the SFL is synonymous with an explicit NULL. As the SFL is the bottom of stack, the IP packet that follows is processed as normal.

If the LSP label is not present due to PHP action in the upstream LSR, two almost equivalent processing actions can take place. Either the SFL can be treated as an LSP label that was not PHPed and then take the additional associated SFL action, which in this case is packet batch counting, or it can be treated as an explicit NULL with associated SFL actions. From the perspective of the measurement system described in this document the behaviour of two approaches are indistinguishable and thus either may be implemented.

5. RFC6374 Packet Loss Measurement

The packet format of an RFC6374 Query message using SFLs is shown in Figure 3.

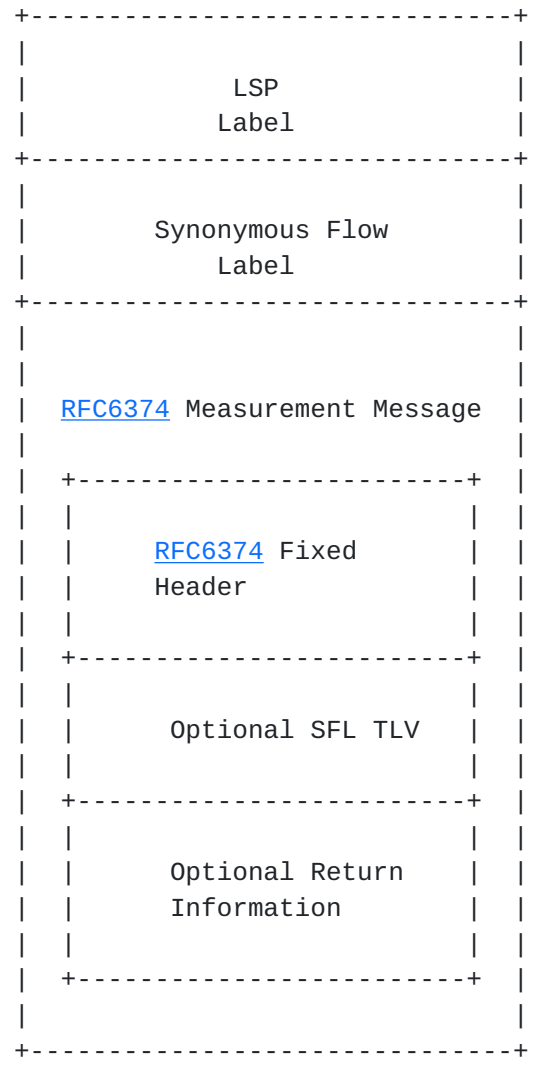

Figure 3: RFC6734 Query Packet with SFL

The MPLS label stack is exactly the same as that used for the user data service packets being instrumented (see Section 4). The RFC6374 measurement message consists of the three components, the RFC6374 fixed header as specified in [RFC6374] carried over the ACH channel type specified the type of measurement being made (currently: loss, delay or loss and delay) as specified in RFC6374.

Two optional TLVs MAY also be carried if needed. The first is the SFL TLV specified in Section 5.1. This is used to provide the implementation with a reminder of the SFL that was used to carry the

RFC6374 message. This is needed because a number of MPLS implementations do not provide the MPLS label stack to the MPLS OAM handler. This TLV is required if RFC6374 messages are sent over UDP (draft-bryant-mpls-RFC63740-over-udp). This TLV MUST be included unless, by some method outside the scope of this document, it is known that this information is not needed by the RFC6374 Responder.

The second set of information that may be needed is the return information that allows the responder send the RFC6374 response to the Querier. This is not needed if the response is requested in-band and the MPLS construct being measured is a point to point LSP, but otherwise MUST be carried. The return address TLV is defined in RFC6378 and the optional UDP Return Object is defined in [I-D.ietf-mpls-rfc6374-udp-return-path].

5.1. SFL TLV

The SFL TLV is shown in Figure 4. This contains the SFL that was carried in the label stack, the FEC that was used to allocate the SFL and the index into the batch of SLs that were allocated for the FEC that corresponds to this SFL.

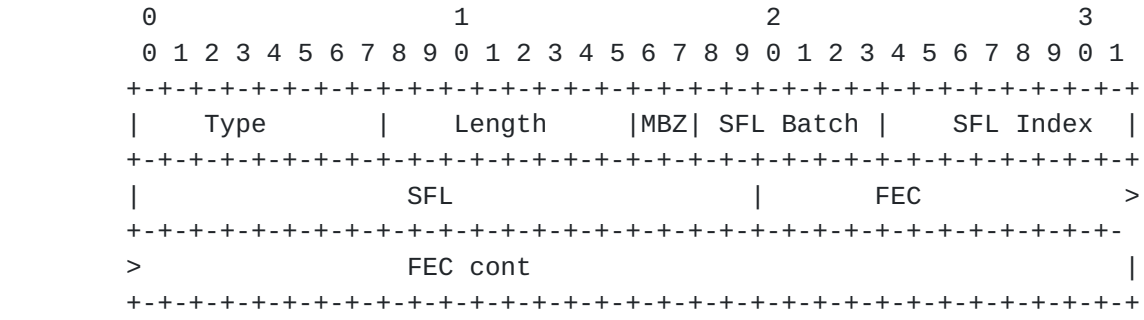

Figure 4: SFL TLV

Where:

Type — Type is set to Synonymous Flow Label (SFL-TLV).

Length — The length of the TLV as specified in [RFC6374].

MBZ — MUST be sent as zero and ignored on receive.

SFL Batch — The SFL batch that this SFL was allocated as part of (see draft-bryant-mpls-sfl-control)

SPL Index The index into the list of SPLs that were assigned against the FEC that corresponds to the SPL.

SFL The SPL used to deliver this packet. This is an MPLS label which is a component of a label stack entry as defined in Section 2.1 of [RFC3032].

FEC The Forwarding Equivalence Class that was used to request this SPL. This is encoded as per Section 3.4.1 of

This information is needed to allow for operation with hardware that discards the MPLS label stack before passing the remainder of the stack to the OAM handler. By providing both the SFL and the FEC plus index into the array of allocated SFLs a number of implementation types are supported.

6. Manageability Considerations

This will be considered in a future version of this document.

7. Privacy Considerations

The inclusion of originating and/or flow information in a packet provides more identity information and hence potentially degrades the privacy of the communication. Whilst the inclusion of the additional granularity does allow greater insight into the flow characteristics it does not specifically identify which node originated the packet other than by inspection of the network at the point of ingress, or inspection of the control protocol packets. This privacy threat may be mitigated by encrypting the control protocol packets, regularly changing the synonymous labels and by concurrently using a number of such labels.

8. Security Considerations

The system described in this memo introduces no additional security vulnerabilities.

9. IANA Considerations

IANA is request to allocate a new TLV from the 0-127 range on the MPLS Loss/Delay Measurement TLV Object Registry:

Type	Description	Reference
TBD	Synonymous Flow Label	This

A value of 4 is recommended.

10. Acknowledgements

TBD

11. References

11.1. Normative References

[I-D.ietf-mpls-rfc6374-udp-return-path]
Bryant, S., Sivabalan, S., and S. Soni, "RFC6374 UDP Return Path", draft-ietf-mpls-rfc6374-udp-return-path-02 (work in progress), 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.

11.2. Informative References

[I-D.bryant-mpls-flow-ident]
Bryant, S. and C. Pignataro, "MPLS Flow Identification", draft-bryant-mpls-flow-ident-00 (work in progress), October 2014.

[I-D.tempia-opsawg-p3m]
Capello, A., Cociglio, M., Castaldelli, L., and A. Bonda, "A packet based method for passive performance monitoring", draft-tempia-opsawg-p3m-04 (work in progress), February 2014.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, September 2011.

Authors' Addresses

Stewart Bryant
Cisco Systems

Email: stbryant@cisco.com

George Swallow
Cisco Systems

Email: swallow@cisco.com

Siva Sivabalan
Cisco Systems

Email: msiva@cisco.com