MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 5, 2021

S. Bryant
Futurewei Technologies Inc
M. Chen
Huawei
G. Swallow
Southend Technical Center
S. Sivabalan
Ciena Corporation
G. Mirsky
ZTE Corp.

October 02, 2020

Synonymous Flow Label Framework
draft-ietf-mpls-sfl-framework-11

Abstract

RFC 8372 (MPLS Flow Identification Considerations) describes the requirement for introducing flow identities within the MPLS architecture. This document describes a method of accomplishing this by using a technique called Synonymous Flow Labels in which labels which mimic the behaviour of other labels provide the identification service. These identifiers can be used to trigger per-flow operations on the packet at the receiving label switching router.

Table of Contents

1. Introduction

[RFC8372] (MPLS Flow Identification Considerations) describes the requirement for introducing flow identities within the MPLS architecture. This document describes a method of providing the required identification by using a technique called Synonymous Flow Labels (SFL) in which labels which mimic the behaviour of other MPLS labels provide the identification service. These identifiers can be used to trigger per-flow operations on the packet at the receiving label switching router.

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 BCP

14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Synonymous Flow Labels

An SFL is defined to be a label that causes exactly the same behaviour at the egress Label Edge Router (LER) as the label it replaces, except that it also causes one or more additional actions that have been previously agreed between the peer LERs to be executed on the packet. There are many possible additional actions such as the measurement of the number of received packets in a flow, triggering an IP Flow Information Export (IPFIX) [RFC7011] capture, triggering other types of Deep Packet Inspection, or identification of the packet source. In, for example, a Performance Monitoring (PM) application, the agreed action could be 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.

To illustrate the use of this technology, we start by considering the case where there is an "application" label in the MPLS label stack. As a first example, let us consider a pseudowire (PW) [RFC3985] on which it is desired to make packet loss measurements. Two labels, synonymous with the PW labels, are obtained from the egress terminating provider edge (T-PE). By alternating between these SFLs and using them in place of the PW label, the PW packets may be batched for counting without any impact on the PW forwarding behavior [RFC8321] (note that strictly only one SFL is needed in this application, but that is an optimization that is a matter for the implementor). The method of obtaining these additional labels is outside the scope of this text, however, one control protocol that provides a method of obtaining SFLs is described in [I-D.bryant-mpls-sfl-control].

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 SFLs source specific, so that batches from one source are marked differently from batches from another source. The sources all operate independently and asynchronously from each other, independently coordinating with the destination. Each ingress LER is thus able to establish its own SFL to identify the sub-flow and thus enable PM per flow.

Finally we need to consider the case where there is no MPLS application label such as occurs when sending IP over an LSP, i.e. there is a single label in the MPLS label stack. In this case introducing an SFL 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, the semantics are similar to an MPLS explicit NULL [RFC3032]. In both of these cases the label has the additional semantics of the SFL.

Note that to achieve the goals set out above, SFLs 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. Application label is present

2. Single label stack

4.1. Application Label Present

Figure 1 shows the case in which both an LSP label and an application label are present in the MPLS label stack. Traffic with no SFL function present runs over the "normal" stack, and SFL-enabled flows run over the SFL stack with the SFL used to indicate the packet batch.

```
 +-----------------+          +-----------------+
 |      LSP        |          |      LSP        |
 |     Label       |          |     Label       |
 |  (May be PHPed) |          |  (May be PHPed) |
 +-----------------+          +-----------------+
 |                 |          |                 |
 |  Application    |          | Synonymous Flow |
 |     Label       |          |     Label       |
 +-----------------+ <= BoS   +-----------------+ <= Bottom of stack
 |                 |          |                 |
 |   Payload       |          |   Payload       |
 |                 |          |                 |
 +-----------------+          +-----------------+


"Normal" Label Stack         Label Stack with SFL
```

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

At the egress LER the LSP label is popped (if present). Then the SFL is processed executing both the synonymous function and the corresponding application function.

4.1.1. Setting TTL and the Traffic Class Bits

The TTL and the Traffic Class bits [RFC5462] in the SFL label stack entry (LSE) would normally be set to the same value as would have been set in the label that the SFL is synonymous with. However, it is recognized that if there is an application need these fields in the SFL Label Stack Entry (LSE) MAY be set these to some other value. An example would be where it was desired to cause the SFL to trigger an action in the TTL expiry exception path as part of the label action.

4.2. Single Label Stack

Figure 2 shows the case in which only an LSP label is present in the MPLS label stack. Traffic with no SFL function present runs over the "normal" stack and SFL-enabled 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 Label Edge Router (LER) to first push the SFL and then to push the LSP label.

```
                             +-----------------+
                             |      LSP        |
                             |     Label       |
                             |  (May be PHPed) |
    +-----------------+      +-----------------+
    |      LSP        |      |                 | <= Synonymous with
    |     Label       |      | Synonymous Flow |    Explicit NULL
    |  (May be PHPed) |      |     Label       |
    +-----------------+ <= BoS  +-----------------+ <= Bottom of stack
    |                 |      |                 |
    |   Payload       |      |   Payload       |
    |                 |      |                 |
    +-----------------+      +-----------------+


 "Normal" Label Stack         Label Stack with SFL
```

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

At the receiving Label Switching Router (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 is 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 processing of the SFL is synonymous with an MPLS 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 the additional associated SFL action is taken when the label is processed. Alternatively, it can be treated as an MPLS Explicit NULL with associated SFL actions. From the perspective of the measurement system described in this document the behaviour of the two approaches is indistinguishable and thus either may be implemented.

4.2.1. Setting TTL and the Traffic Class Bits

The TTL and the Traffic Class considerations described in Section 4.1.1 apply.

4.3. Aggregation of SFL Actions

There are cases where it is desirable to aggregate an SFL action against a number of labels. For example, where it is desirable to have one counter record the number of packets received over a group of application labels, or where the number of labels used by a single application is large, and the resultant increase in the number of allocated labels needed to support the SFL actions may becomes too large to be viable. In these circumstances it would be necessary to introduce an additional label in the stack to act as an aggregate instruction. This is not strictly a synonymous action in that the SFL is not replacing an existing label, but is somewhat similar to the single label case shown in Section 4.2, and the same signalling, management and configuration tools would be applicable.

```
                                 +-----------------+
                                 |      LSP        |
                                 |     Label       |
                                 |  (May be PHPed) |
    +-----------------+          +-----------------+
    |      LSP        |          |                 |
    |     Label       |          |   Aggregate     |
    |  (May be PHPed) |          |      SFL        |
    +-----------------+          +-----------------+
    |                 |          |                 |
    |  Application    |          |  Application    |
    |     Label       |          |     Label       |
    +-----------------+ <=BoS    +-----------------+ <= Bottom of stack
    |                 |          |                 |
    |   Payload       |          |   Payload       |
    |                 |          |                 |
    +-----------------+          +-----------------+


   "Normal" Label Stack         Label Stack with SFL
```

Figure 3: Aggregate SFL Actions

The Aggregate SFL is shown in the label stack depicted in Figure 3 as preceding the application label, however the choice of position before, or after, the application label will be application specific. In the case described in Section 4.1, by definition the SFL has the full application context. In this case the positioning will depend on whether the SFL action needs the full context of the application to perform its action and whether the complexity of the application will be increased by finding an SFL following the application label.

5. Equal Cost Multipath Considerations

The introduction of an SFL to an existing flow may cause that flow to take a different path through the network under conditions of Equal Cost Multi-path (ECMP). This in turn may invalidate certain uses of the SFL such as performance measurement applications. Where this is a problem there are two solutions worthy of consideration:

1. The operator MAY elect to always run with the SFL in place in the MPLS label stack.

2. The operator can elect to use [RFC6790] Entropy Labels in a network that fully supports this type of ECMP. If this approach is adopted, the intervening MPLS network MUST NOT load balance on any packet field other than the entropy label. Note that this is stricter than the text in Section 4.3 of [RFC6790].

6. Privacy Considerations

IETF concerns on pervasive monitoring are described in [RFC7258]. The inclusion of originating and/or flow information in a packet provides more identity information and hence potentially degrades the privacy of the communication to an attacker in a position to observe the added identifier. 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 unless the attacker can inspect 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, by regularly changing the synonymous labels or by concurrently using a number of such labels, including the use of a combination of those methods. Minimizing the scope of the identity indication can be useful in minimizing the observability of the flow characteristics. Whenever IPFIX or other DPI technique is used, their relavent privacy considerations apply.

7. Security Considerations

There are no new security issues associated with the MPLS data plane. Any control protocol used to request SFLs will need to ensure the legitimacy of the request, i.e. that the requesting node is authorized to make that SFL request by the network operator.

8. IANA Considerations

This draft makes no IANA requests.

9. Contributing Authors

Zhenbin Li
Huawei
Email: lizhenbin@huawei.com

10. References

10.1. 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>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[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>.

10.2. Informative References

[I-D.bryant-mpls-sfl-control] Bryant, S., Swallow, G., and S. Sivabalan, "A Simple Control Protocol for MPLS SFLs", draft-bryant-mpls-sfl-control-08 (work in progress), June 2020.

[RFC3985] Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, DOI 10.17487/RFC3985, March 2005, <https://www.rfc-editor.org/info/rfc3985>.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an Attack", BCP 188, RFC 7258, DOI 10.17487/RFC7258, May 2014, <https://www.rfc-editor.org/info/rfc7258>.

[RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

[RFC8372] Bryant, S., Pignataro, C., Chen, M., Li, Z., and G. Mirsky, "MPLS Flow Identification Considerations", RFC 8372, DOI 10.17487/RFC8372, May 2018, <https://www.rfc-editor.org/info/rfc8372>.

Authors' Addresses

Stewart Bryant
Futurewei Technologies Inc

Email: sb@stewartbryant.com

Mach Chen
Huawei

Email: mach.chen@huawei.com

George Swallow
Southend Technical Center

Email: swallow.ietf@gmail.com

Siva Sivabalan
Ciena Corporation

Email: ssivabal@ciena.com

Gregory Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com