Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 16, 2015

M. Chen
X. Xu
Z. Li
Huawei
L. Fang
Microsoft
G. Mirsky
Ericsson

October 13, 2014

MultiProtocol Label Switching (MPLS) Source Label
draft-chen-mpls-source-label-06

Abstract

A MultiProtocol Label Switching (MPLS) label was originally defined to identify a Forwarding Equivalence Class (FEC). A packet is assigned to a specific FEC based on its network layer destination address, and optionally Class of Service. It’s difficult or even impossible to derive the source identity information from the label. For some applications, source identification is a critical requirement. For example, performance monitoring, where the monitoring node needs to identify where a packet was sent from.

This document introduces the concept of Source Identifier (SI) that identifies the ingress Label Switching Router (LSR) of a Label Switched Path (LSP). A SI is unique within a domain that is referred to as Source Identifier Administrative Domain (SIAD).

This document also introduces the concept of Source Label (SL) that is carried in the label stack and carries the SI of the ingress LSR of an LSP. Source Label is preceded by a Source Label Indicator (SLI) when included the label stack and is not used for forwarding.

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 RFC 2119 [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. Problem Statement and Introduction

A MultiProtocol Label Switching (MPLS) label [RFC3031] was originally defined for packet forwarding and assumes the forwarding/destination address semantics. As no source identity information is carried in the label stack, in many cases there is no way to directly derive the source identity information from the label or label stack.

MPLS LSPs can be categorized into four different types:

- Point-to-Point (P2P)

- Point-to-Multipoint (P2MP)

- Multipoint-to-Point (MP2P)

- Multipoint-to-Multipoint (MP2MP)

For P2P and P2MP LSPs (e.g., the Resource Reservation Protocol Traffic Engineering (RSVP-TE) [RFC3209] based and statically configured P2P and P2MP LSPs), the source identity may be implicitly derived by the egress LSR from the label when Penultimate Hop Popping (PHP) is disabled and the correlation between ingress LSR and the LSP is explicitly signaled through the control plane. Such LSP may be characterized as MPLS-TP LSP [RFC5960].

However, for MP2P and MP2MP LSPs (e.g., the Label Distribution Protocol (LDP) based LSPs [RFC5036] [RFC6388], and Layer 3 Private Network (L3VPN) [RFC4364] LSPs), ingress LSRs of those LSPs cannot be identified by egress LSRs.

Comparing to the pure IP forwarding where both source and destination addresses are encoded in the IP packet header, the essential issue of the MPLS encoding is that the label stack does not explicitly include any source identity information. For some applications, source identification is a critical requirement. For example, performance monitoring, the monitoring nodes need to identify where packets were sent from and then can count the packets according to some constraints.

This document introduces the concept of Source Label (SL). An SL is carried in the label stack and carries the identifier of the ingress LSR that originated the MPLS frame.

2. Terminology

SI - Source Identifier

SIAD - Source Identifier Administrative Domain

SL - Source Label

SLC - Source Label Capability

SLI - Source Label Indicator

3. Source Label

A Source Label is defined to carry an identifier (Source Identifier) of a node that is (one of) the ingress LSR(s) to specific LSP. Source Label SHOULD NOT be used for forwarding and is not signaled.

A Source Identifier (SI) is a number in the range of [16, 65535]. Each node in a domain MUST be allocated one or more unique SIs, the domain is referred as a "Source Identifier Administrative Domain" (SIAD). For most of the use cases, one SI per LSR would be sufficient. But for some cases, there may be need for more than one SIs. For example, in the L3VPN scenario, it may be necessary to allocate a dedicated SI to identify each VPN instance.

In order to indicate whether a label is a Source Label, a Source Label Indicator (SLI) is introduced. The SLI is a special purpose label [RFC7274] that is placed immediately before the source label in the label stack, which is used to indicate that the next label in the label stack is the Source Label. The value of SLI is TBD1. The SL is an example of context label [RFC5331], the SLI is the context.

To prevent the Source Label from leaking to unintended domains, two aspects need to be considered:

- In the control plane, the Source Label MUST NOT be distributed outside the SIAD where it is used. Since the ingress LSR is based on the Source Label Capability signaled by the egress LSR to determine whether to insert the Source Label, the SLC signaling MUST make sure that the SLC will not be signaled to the LSRs that reside in other SIADs.

- In the data plane, the domain boundary nodes (e.g., the ASBR) SHOULD have the capability to filter out the packets that carry the SL/SLI and are received from other SIADs. For example, some policies (e.g., using ACL) could be deployed at the ASBR to filter out the packets that carry SL/SLI and are from other SIADs.

4. Performance Measurement Use Case

There are two general types of performance measurement: one is active performance measurement, and the other is passive performance measurement.

In active performance measurement the receiver measures the injected packets to evaluate the performance of a path. The active measurement measures the performance of the extra injected packets. The IP Performance Metrics (IPPM) working group has defined specifications [RFC4656][RFC5357] for active performance measurement.

In passive performance measurement, no additional traffic is injected into the flow and measurements are taken to record the performance metrics of the data traffic. The MPLS performance measurement protocol [RFC6374] for packet loss is an example of passive performance measurement, but currently it can only be measured for MPLS-TE LSPs. For a specific receiver, in order to count the received packets of a flow, the system doing the measurement (e.g., egress router) needs to know which target flow a received packet belongs to. Source identification is therefore necessary. Source identification may be achieved by including appropriate MEP-ID [RFC6428].

As discussed in the previous section, the existing MPLS label or label stack does not carry the source information. So, for an LSP, the ingress LSR can put its SI in the Source Label, and then the egress LSR can use the SI to identify the packet’s source, in order to facilitate accounting.

5. Data Plane Processing

5.1. Ingress LSR

For an LSP, the ingress LSR MUST make sure that the egress LSR is able to process the Source Label before inserting the SLI/SL combination into the label stack. Therefore, an egress LSR SHOULD signal (see Section 6) to the ingress LSR whether it is able to process the Source Label. Once the ingress LSR knows that the egress LSR can process Source Label, it can choose whether or not to insert the SL and SLI into the label stack.

When an SL to be included in a label stack, the steps are as follows:

1. Push the SL, the TTL of the SL MUST be set to 1, the BoS bit for the SL depends on whether the SL is the bottom label. Setting and interpretation of TC field of the SL is for further study;

2. Push the SLI, the TTL and TC fields for the SLI MUST be set to the same values as for the LSP Label (L);

3. Push the LSP Label (L).

Then the label stack looks like: <...L, SLI, SL [,...]>. There MAY be multiple combinations of SLI and SL inserted into the label stack, each combination is related to an LSP. For the given LSP, only one combination of SLI and SL MUST be inserted.

5.2. Transit LSR

There is no change in forwarding behavior for transit LSRs. If a transit LSR can recognize the SLI, it can use the SL to collect traffic throughput and/or measure the performance of the LSP.

5.3. Egress LSR

When an egress LSR receives a packet with a SLI/SL combination, if the egress LSR is able to process the SL; it pops the LSP label (if any), SLI and SL; then processes remaining packet header as normal. If the egress LSR is not able to process the SLI, the packet SHOULD be dropped as specified for the handling of any unknown label according to [RFC3031].

5.4. Penultimate Hop LSR

There is no change in forwarding behavior for the penultimate hop LSR.

6. Source Label Capability Signaling

Before inserting a Source Label in the label stack, an ingress LSR SHOULD know whether the egress LSR is able to process the SLI and SL. Therefore, an egress LSR SHOULD signal to the ingress LSRs its ability to process the SLI and SL. This is called Source Label Capability (SLC), it is very similar to the "Entropy Label Capability (ELC)"[RFC6790].

6.1. LDP Extensions

A new LDP TLV [RFC5036], SLC TLV, is defined to signal an egress’s ability to process Source Label. The SLC TLV MAY appear as an Optional Parameter of the Label Mapping Message. The presence of the SLC TLV in a Label Mapping Message indicates to ingress LSRs that the egress LSR can process Source Labels for the associated LSP.

The structure of the SLC TLV is shown below.

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |U|F|       Type (TBD2)        |           Length (0)           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Source Label Capability TLV

This U bit MUST be set to 1. If the SLC TLV is not understood by the receiver, then it MUST be ignored.

This F bit MUST be set to 1. Since the SLC TLV is going to be propagated hop-by-hop, it should be forwarded even by nodes that may not understand it.

Type: TBD2.

Length field: This field specifies the total length in octets of the SLC TLV and is defined to be 0.

An LSR that receives a Label Mapping with the SLC TLV but does not understand it MUST propagate it intact to its neighbors and MUST NOT send a notification to the sender (following the meaning of the U- and F-bits). If the LSR has no other neighbors and does not understand the SLC TLV, means it is the ingress LSR, it could just ignore it. An LSR X may receive multiple Label Mappings for a given FEC F from its neighbors. In its turn, X may advertise a Label Mapping for F to its neighbors. If X understands the SLC TLV, and if any of the advertisements it received for FEC F does not include the SLC TLV, X MUST NOT include the SLC TLV in its own advertisements of F. If all the advertised Mappings for F include the SLC TLV, then X MUST advertise its Mapping for F with the SLC TLV. If any of X’s neighbors resends its Mapping, sends a new Mapping or sends a Label Withdraw for a previously advertised Mapping for F, X MUST re-evaluate the status of SLC for FEC F, and, if there is a change, X MUST re-advertise its Mapping for F with the updated status of SLC.

LDP is normally running within an AS, technically, it can be deployed across ASes. An implementation supports the SLC MUST support a per-session/per-interface configuration item to enable/disable the SLC. For the session/interface that connects to other SLADs, the SLC MUST be disabled.

6.2. BGP Extensions

When Border Gateway Protocol (BGP) [RFC4271] is used for distributing Network Layer Reachability Information (NLRI) as described in, for example, [RFC3107], [RFC4364], the BGP UPDATE message may include the

SLC attribute as part of the Path Attributes. This is an optional, non-transitive BGP attribute of value TBD3. The inclusion of this attribute with an NLRI indicates that the advertising BGP router can process Source Labels as an egress LSR for all routes in that NLRI.

A BGP speaker S that originates an UPDATE should include the SLC attribute only if both of the following are true:

A1: S sets the BGP NEXT_HOP attribute to itself AND

A2: S can process source labels.

Suppose a BGP speaker T receives an UPDATE U with the SLC attribute. T has two choices. T can simply re-advertise U with the SLC attribute if either of the following is true:

B1: T does not change the NEXT_HOP attribute OR

B2: T simply swaps labels without popping the entire label stack and processing the payload below.

An example of the use of B1 is Route Reflectors. However, if T changes the NEXT_HOP attribute for U and in the data plane pops the entire label stack to process the payload, T MAY include an SLC attribute for UPDATE U’ if both of the following are true:

C1: T sets the NEXT_HOP attribute of U’ to itself AND

C2: T can process source labels. Otherwise, T MUST remove the SLC attribute.

6.2.1. Sending/Receiving Restriction

An implementation that supports the SLC MUST support per-session configuration item, SL_SESSION, that indicates whether the SLC is enabled or disabled for use on that session.

- The default value of SL_SESSION, for EBGP sessions, MUST be "disabled".

- The default value of SL_SESSION, for IBGP and confederation-EBGP [RFC5065]sessions, SHOULD be "enabled."

The SLC attribute MUST NOT be sent on any BGP session for which SL_SESSION is disabled.

If an SLC attribute is received on a BGP session for which SL_SESSION is disabled, the attribute MUST be treated exactly as if it were an unrecognized non-transitive attribute. That is, "it MUST be quietly ignored and not passed along to other BGP peers" (see [RFC4271], section 5).

6.3. IGP Extensions

IGP based SLC applies to the scenarios where IGP is used for label mapping (e.g., Segment Routing). IGP SLC signaling is defined in [I-D.chen-isis-source-identifier-distribution] and [I-D.chen-ospf-source-identifier-distribution], the presence of a Source Identifier TLV/sub-TLV MUST be interpreted as support of SLC by the LSR. That means the SLC is implicitly indicated by receiving a SI distribution from an LSR.

7. Source Identifier Distribution

Based on the Source Identifier, an egress or intermediate LSR can identify from where an MPLS packet is sent. To achieve this, the egress and/or intermediate LSRs have to know which ingress LSR is related to which Source Identifier before using the Source Identifier to derive the source information. Therefore, there needs to be a mechanism to distribute the mapping information between an ingress LSR and its SI(s).

IGP based SI distribution documents, [I-D.chen-isis-source-identifier-distribution], [I-D.chen-ospf-source-identifier-distribution], define extensions to corresponding IGP protocols necessary for intra-AS scenario.

For inter-AS scenario, BGP extension is a naturally choice and can be used to convey SI mapping information from one AS to other ASes. The BGP extension draft is work in progress. For BGP based SI distribution, it requires that SIs MUST not be sent out of a SIAD. The similar sending and receiving restriction as defined in Section 6.2.1 is also needed.

8. IANA Considerations

8.1. Source Label Indication

IANA is required to allocate a special purpose label (TBD1) for the Source Label Indicator (SLI) from the "Multiprotocol Label Switching Architecture (MPLS) Label Values" Registry.

8.2. LDP Source Label Capability TLV

IANA is requested to allocate a value of TBD2 from the IETF Consensus range (0x0001-0x07FF) in the "TLV Type Name Space" registry as the "Source Label Capability TLV".

8.3. BGP Source Label Capability Attribute

IANA is requested to allocate a Path Attribute Type Code TBD3 from the "BGP Path Attributes" registry as the "BGP Source Label Capability Attribute".

9. Security Considerations

This document introduces the SIAD that is the scope of a SL. The SLC and SI MUST NOT be signaled and distributed outside one SIAD. BGP based SLC and SI distribution is controlled by SL_SESSION configuration. Improper configuration on both ends of an EBGP connection could result in the SLC and SI being passed from one SIAD to another. This would likely result in potential SI conflicts.

To prevent packets carrying SL/SLI from leaking from one SIAD to another, the SIAD boundary nodes SHOULD deploy some policies (e.g., ACL) to filter out the packets. Specifically, in the sending end, the SIAD boundary node SHOULD filter out the packets that carry the SLI and are sent to other SIADs; in the receiving end, the SIAD boundary node SHOULD drop the packets that carry the SLI and are from other SIADs.

10. Acknowledgements

The process of "Source Label Capability Signaling" is largely referred to the process of "ELC signaling"[RFC6790].

The authors would like to thank Carlos Pignataro, Loa Andersson , Curtis Villamizar, Eric Osborne, Eric Rosen, Yimin Shen, Lizhong Jin, Ross Callon and Yakov Rekhter for their review, suggestion and comments to this document.

11. References

11.1. Normative References

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

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[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.

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

[RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.

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

[RFC7274] Kompella, K., Andersson, L., and A. Farrel, "Allocating and Retiring Special-Purpose MPLS Labels", RFC 7274, June 2014.

11.2. Informative References

[I-D.chen-isis-source-identifier-distribution]
Chen, M. and G. Mirsky, "Extensions to ISIS for Source Identifier Distribution", draft-chen-isis-source-identifier-distribution-00 (work in progress), October 2014.

[I-D.chen-ospf-source-identifier-distribution]
Chen, M. and G. Mirsky, "Extensions to OSPF for Source Identifier Distribution", draft-chen-ospf-source-identifier-distribution-00 (work in progress), October 2014.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

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

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, September 2006.

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

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, August 2007.

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

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, October 2008.

[RFC5960] Frost, D., Bryant, S., and M. Bocci, "MPLS Transport Profile Data Plane Architecture", RFC 5960, August 2010.

[RFC6388] Wijnands, IJ., Minei, I., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, November 2011.

[RFC6428] Allan, D., Swallow Ed. , G., and J. Drake Ed. , "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428, November 2011.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, November 2012.

Authors’ Addresses

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Xiaohu Xu
Huawei

Email: xuxiaohu@huawei.com

Zhenbin Li
Huawei

Email: lizhenbin@huawei.com

Luyuan Fang
Microsoft

Email: lufang@microsoft.com

Greg Mirsky
Ericsson

Email: Gregory.mirsky@ericsson.com