Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

S. Kini, Ed.
Ericsson
H. Gredler
Juniper Networks

October 21, 2013

Detecting Multi-Protocol Label Switching (MPLS) Data Plane Failures in Source Routed LSPs
draft-kini-spring-mpls-lsp-ping-00

Abstract

MPLS has defined mechanisms for fault detection and isolation and mechanisms for reliably sending an echo reply in RFC 4379. Source routed MPLS LSPs are a technique being proposed to address new use-cases. This document describes how mechanisms defined for MPLS fault detection and isolation can be applied for source routed LSPs.

Table of Contents

1. Introduction

Multi Protocol Label Switching (MPLS) has defined in [RFC4379] a simple and efficient mechanism to detect data plane failures in Label Switched Paths (LSP) by specifying information to be carried in an MPLS "echo request" and "echo reply" for the purposes of fault detection and isolation, and mechanisms for reliably sending the echo reply. The functionality is modeled after the ping/traceroute paradigm (ICMP echo request [RFC0792]) and is typically referred to as MPLS-ping and MPLS-traceroute.

Source routed LSP is a technique by which the ingress stacks a set of tunnels to route the packet through an explicit-route. Newer use-cases (e.g. [OAM-UC], [I-D.geib-spring-oam-usecase]) are being explored using this technique and detecting data plane failures is a basic requirement in all of them. This document describes how the procedures defined in [RFC4379] can be applied to a source routed LSP.

1.1. 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].

2. Abbreviations and Terminology

TTL - Time to Live

OAM - Operation, Administration, Management/Maintenance

LSP - Label Switched Path

FEC - Forwarding Equivalence Class

SPRING - Source Packet Routing in Networking

3. Service Labels

One of the proposals for source routed LSPs is to include service labels in the MPLS label stack. These service labels are used to apply a service (as indicated by the service label) to the packet at the intermediate LSRs along the explicit-route. Since these labels are part of the MPLS label stack these have implications on MPLS OAM. This document describes how the procedures of [RFC4379] can be applied to in the absence of service-labels in Section 5. Additional considerations for service labels are included in Section 6 and requires further discussion.

4. Packet Format

4.1. Target FEC Stack

The following new FEC Type sub-TLVs are defined to accommodate the distribution of labels by Interior Gateway Protocols (IGP) OSPF and ISIS ([I-D.gredler-rtgwg-igp-label-advertisement], [I-D.gredler-isis-label-advertisement], [I-D.gredler-ospf-label-advertisement], [I-D.previdi-isis-segment-routing-extensions], [I-D.psenak-ospf-segment-routing-extensions]).

Sub-Type(suggested values)	Length	Value Field
17	5	OSPF IPv4 Prefix
18	17	OSPF IPv6 Prefix
19	5	ISIS IPv4 Prefix
20	17	ISIS IPv6 Prefix

Table 1

4.2. OSPF IPv4 Prefix

This value of this sub-TLV is encoded the same as the "Generic IPv4 Prefix" defined in section 3.2.13 of [RFC4379].

4.3. OSPF IPv6 Prefix

This value of this sub-TLV is encoded the same as the "Generic IPv6 Prefix" defined in section 3.2.14 of [RFC4379].

4.4. ISIS IPv4 Prefix

This value of this sub-TLV is encoded the same as the "Generic IPv4 Prefix" defined in section 3.2.13 of [RFC4379].

4.5. ISIS IPv6 Prefix

This value of this sub-TLV is encoded the same as the "Generic IPv6 Prefix" defined in section 3.2.14 of [RFC4379].

5. MPLS ping and trace of a source routed LSP

The MPLS ping procedures described in [RFC4379] can be applied unchanged to a source routed LSP. The ingress should encapsulate the "echo request" with the label stack just as any data packet and send it on the source routed LSP. Sometimes it is useful to ping a specific tunnel that is used in a source routed LSP. In this case the entire label stack of the source routed LSP must be used, but the TTL of labels below the label of the tunnel that being debugged must be set to zero.

When tracing a LSP according to the procedures in [RFC4379] the TTL is incremented by one in order to trace the path sequentially along the LSP. However when a source routed LSP has to be traced there are as many TTLs as there are labels in the stack. The LSR that initiates the traceroute SHOULD start by setting the TTL to 1 for the tunnel in the LSP's label stack it wants to start the tracing from, the TTL of all outer labels in the stack to the max value, and the TTL of all the inner labels in the stack to zero. Thus a typical start to the traceroute would have a TTL of 1 for the outermost label and all the inner labels would have TTL 0. If the FEC Stack TLV is included it should contain only those for the inner stacked tunnels. The lack of an echo response or the Return Code/Subcode should be used to diagnose the tunnel as described in [RFC4379]. When the tracing of a tunnel in the stack is complete, then the next tunnel in the stack should be traced. The end of a tunnel can be detected from

the "Return Code" when it indicates that the responding LSR is an egress for the stack at depth 1. Thus the traceroute procedures in [RFC4379] can be recursively applied to traceroute a source routed LSP.

6. Issues with non-forwarding labels

Source stacking can be optionally used to apply services on the packet at a LSR along the path, where a label in the stack is used to trigger service application. A data plane failure detection and isolation mechanism should provide its functionality without applying these services. This is mandatory for services that are stateful, though for stateless services [RFC4379] could be used as-is. It MAY also provide a mechanism to detect and isolate faults within the service function itself.

To prevent services from being applied to an "echo request" packet, the TTL of service labels MUST be 0. However TTL processing rules of a service label must be the same as any MPLS label. Due to this a TTL of 0 in the service label would prevent the packet from being forwarded beyond the LSR that provides the service. To avoid this problem, the originator of the "echo request" must remove those service labels from the stack upto the tunnel that is being currently traced. In other words the ingress must remove all service-labels above the label of the tunnel being currently traced, but retain service labels below it when sending the echo request. Note that load balancing may affect the path when the service labels are removed, resulting in a newer path being traversed. However this new path is potentially different only upto the LSR that provides the service. Since this portion of the path was traced when the tunnels above this tunnel in the stack were traced and followed the exact path as the source routed LSP, this should not be a major concern. Sometimes the newer path may have a problem that was not in the original path resulting in a false positive. In such a case the original path can be traversed by changing the label stack to reach the intermediate LSR with labels that route along each hop explicitly.

7. Acknowledgements

The authors would like to thank TBD for their comments.

8. IANA Considerations

New Sub-Types for the FEC Stack TLV are required to be allocated.

9. Security Considerations

10. References

10.1. Normative References

[I-D.gredler-isis-label-advertisement]
Gredler, H., Amante, S., Scholl, T., and L. Jalil, "Advertising MPLS labels in IS-IS", draft-gredler-isis-label-advertisement-03 (work in progress), May 2013.

[I-D.gredler-ospf-label-advertisement]
Gredler, H., Amante, S., Scholl, T., and L. Jalil, "Advertising MPLS labels in OSPF", draft-gredler-ospf-label-advertisement-03 (work in progress), May 2013.

[I-D.gredler-rtgwg-igp-label-advertisement]
Gredler, H., Amante, S., Scholl, T., and L. Jalil, "Advertising MPLS labels in IGPs", draft-gredler-rtgwg-igp-label-advertisement-05 (work in progress), May 2013.

[I-D.previdi-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., and S. Litkowski, "IS-IS Extensions for Segment Routing", draft-previdi-isis-segment-routing-extensions-03 (work in progress), October 2013.

[I-D.psenak-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., and W. Henderickx, "OSPF Extensions for Segment Routing", draft-psenak-ospf-segment-routing-extensions-03 (work in progress), October 2013.

[OAM-UC] Google, "Google Blackbox Monitoring", 2012, <https://ripe65.ripe.net/presentations/828-RIPE65.Talk29.Google_Blackbox_Monitoring.pdf>.

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

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

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

[RFC6424] Bahadur, N., Kompella, K., and G. Swallow, "Mechanism for Performing Label Switched Path Ping (LSP Ping) over MPLS Tunnels", RFC 6424, November 2011.

10.2. Informative References

[I-D.filsfils-rtgwg-segment-routing-use-cases]
Filsfils, C., Francois, P., Previdi, S., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., Kini, S., and E. Crabbe, "Segment Routing Use Cases", draft-filsfils-rtgwg-segment-routing-use-cases-02 (work in progress), October 2013.

[I-D.filsfils-rtgwg-segment-routing]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Architecture", draft-filsfils-rtgwg-segment-routing-01 (work in progress), October 2013.

[I-D.geib-spring-oam-usecase]
Geib, R., "Use case for a scalable and topology aware MPLS data plane monitoring system", draft-geib-spring-oam-usecase-00 (work in progress), October 2013.

Authors' Addresses

Sriganesh Kini (editor)
Ericsson

Email: sriganesh.kini@ericsson.com

Hannes Gredler
Juniper Networks

Email: hannes@juniper.net