

spring
Internet-Draft
Intended status: Informational
Expires: April 23, 2017

R. Geib, Ed.
Deutsche Telekom
C. Filsfils
C. Pignataro, Ed.
N. Kumar
Cisco
October 20, 2016

A Scalable and Topology-Aware MPLS Dataplane Monitoring System draft-ietf-spring-oam-usecase-04

Abstract

This document describes features of a path monitoring system and related use cases. Segment based routing enables a scalable and simple method to monitor data plane liveliness of the complete set of paths belonging to a single domain. The MPLS monitoring system adds features to the traditional MPLS ping and LSP trace, in a very complementary way. MPLS topology awareness reduces management and control plane involvement of OAM measurements while enabling new OAM features.

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 April 23, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Acronyms	2
2.	Introduction	2
3.	An MPLS Topology-Aware Path Monitoring System	4
4.	SR-based Path Monitoring Use Case Illustration	6
4.1.	Use Case 1 - LSP Dataplane Monitoring	6
4.2.	Use Case 2 - Monitoring a Remote Bundle	8
4.3.	Use Case 3 - Fault Localization	9
5.	Failure Notification from PMS to LERi	9
6.	Applying SR to Monitoring non-SR based LSPs (LDP and possibly RSVP-TE)	9
7.	PMS Monitoring of Different Segment ID Types	10
8.	Connectivity Verification Using PMS	10
9.	Extensions of Specifications Relevant to this Use Case	10
10.	IANA Considerations	10
11.	Security Considerations	10
12.	Acknowledgements	11
13.	References	11
13.1.	Normative References	11
13.2.	Informative References	11
	Authors' Addresses	12

[1.](#) Acronyms

ECMP Equal-Cost Multi-Path
 IGP Interior Gateway Protocol
 LER Label Edge Router
 LSP Label Switched Path
 LSR Label Switching Router
 OAM Operations, Administration, and Maintenance
 PMS Path Monitoring System
 SID Segment Identifier
 SR Segment Routing
 SRGB Segment Routing Global Block

[2.](#) Introduction

It is essential for a network operator to monitor all the forwarding paths observed by the transported user packets. The monitoring packet is expected to be forwarded in dataplane in a similar way as

user packets. Segment Routing enables forwarding of packets along pre-defined paths and segments and thus a Segment Routed monitoring packet can stay in dataplane while passing along one or more segments to be monitored.

This document describes a system using MPLS data plane path monitoring capabilities. The use cases introduced here is limited to a single IGP MPLS domain.

The system applies to monitoring of LDP LSP's as well as to monitoring of Segment Routed LSP's. As compared to LDP, Segment Routing is expected to simplify the system by enabling MPLS topology detection based on IGP signaled segments as specified at [[I-D.ietf-isis-segment-routing-extensions](#)] and [[I-D.ietf-ospf-segment-routing-extensions](#)]. Thus a centralised and MPLS topology aware monitoring unit can be realized in a Segment Routed domain. This topology awareness can be used for OAM purposes as described by this document.

The MPLS path monitoring system described by this document can be realised with pre-Segment Routing (SR) based technology. Making such a pre-SR MPLS monitoring system aware of a domain's complete MPLS topology requires e.g. management plane access. To avoid the use of stale MPLS label information, IGP must be monitored and MPLS topology must be timely aligned with IGP topology. Obviously, enhancing IGPs to exchange of MPLS topology information as done by SR significantly simplifies and stabilises such an MPLS path monitoring system.

This document adopts the terminology and framework described in [[I-D.ietf-spring-segment-routing](#)].

The system offers several benefits for network monitoring. A single centralized monitoring device is able to monitor the complete set of a domain's forwarding paths. Monitoring packets never leave data plane. MPLS path trace function (whose specification and features are not part of this use case) is required, if the actual data plane of a router should be checked against its control plane. SR capabilities allow to direct MPLS OAM packets from a centralized monitoring system to any router within a domain whose path should be traced.

In addition to monitoring paths, problem localization is required. Faults can be localized:

- o by IGP LSA analysis.
- o correlation between different SR based monitoring probes.

- o by any MPLS traceroute method (possibly in combination with SR based path stacks).

Topology awareness is an essential part of link state IGPs. Adding MPLS topology awareness to an IGP speaking device hence enables a simple and scalable data plane based monitoring mechanism.

MPLS OAM offers flexible features to recognise and execute data paths of an MPLS domain. By utilising the ECMP related tool set offered e.g. by [RFC 4379](#) [[RFC4379](#)], a segment based routing LSP monitoring system may:

- o easily detect ECMP functionality and properties of paths at data level.
- o construct monitoring packets executing desired paths also if ECMP is present.
- o limit the MPLS label stack of an OAM packet to a minimum of 3 labels.

Alternatively, any path may be executed by building suitable label stacks. This allows path execution without ECMP awareness.

The MPLS Path Monitoring System (PMS) may be any server residing at a single interface of the domain to be monitored. It doesn't have to support any specialised protocol stack, it just should be capable of understanding the topology and building the monitoring probe packet with the right segment stack. The monitoring probe packet could be BFD or LSP Ping packet or any other OAM format that PMS supports. As long as the monitoring packet returns back to the server, the path can be considered as validated. The MPLS monitoring servers are the single entities pushing monitoring packet label stacks. If the depth of label stacks to be pushed by a path monitoring system (PMS) are of concern for a domain, a dedicated server based path monitoring architecture allows limiting monitoring related label stack pushes to these servers.

Documents discussing SR OAM requirements and possible solutions to allow SR usage as described by this document have been submitted already, see [[I-D.ietf-spring-sr-oam-requirement](#)] and [[I-D.ietf-mpls-spring-lsp-ping](#)].

3. An MPLS Topology-Aware Path Monitoring System

An MPLS PMS which is able to learn the IGP LSDB (including the SID's) is able to execute arbitrary chains of label switched paths. It can send pure monitoring packets along such a path chain or it can direct

suitable MPLS OAM packets to any node along a path segment. Segment Routing here is used as a means of adding label stacks and hence transport to standard MPLS OAM packets, which then detect correspondence of control and data plane of this (or any other addressed) path. Any node connected to an SR domain is MPLS topology aware (the node knows all related IP addresses, SR SIDs and MPLS labels). Thus a PMS connected to an MPLS SR domain just needs to set up a topology data base for monitoring purposes.

Let us describe how the PMS constructs a labels stack to transport a packet to LER i, monitor its path to LER j and then receive the packet back.

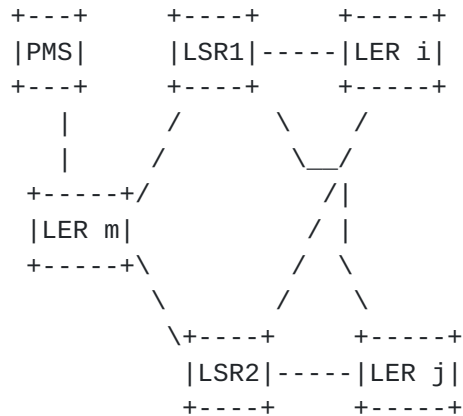
The PMS may do so by sending packets carrying the following MPLS label stack information:

- o Top Label: a path from PMS to LER i, which is expressed as Node SID of LER i.
- o Next Label: the path that needs to be monitored from LER i to LER j. If this path is a single physical interface (or a bundle of connected interfaces), it can be expressed by the related AdjSID. If the shortest path from LER i to LER j is supposed to be monitored, the Node-SID (LER j) can be used. Another option is to insert a list of segments expressing the desired path (hop by hop as an extreme case). If LER i pushes a stack of Labels based on a SR policy decision and this stack of LSPs is to be monitored, the PMS needs an interface to collect the information enabling it to address this SR created path.
- o Next Label or address: the path back to the PMS. Likely, no further segment/label is required here. Indeed, once the packet reaches LER j, the 'steering' part of the solution is done and the probe just needs to return to the PMS. This is best achieved by popping the MPLS stack and revealing a probe packet with PMS as destination address (note that in this case, the source and destination addresses could be the same). If an IP address is applied, no SID/label has to be assigned to the PMS (if it is a host/server residing in an IP subnet outside the MPLS domain).

Note: a deployment might prefer not to connect the PMS to the MPLS domain. if the PMS is an IP host not connected to the MPLS domain, the PMS can send its probe with the list of SIDs/Labels onto a suitable tunnel providing an MPLS access to a router which is part of the monitored MPLS domain.

4. SR-based Path Monitoring Use Case Illustration

4.1. Use Case 1 - LSP Dataplane Monitoring



Example of a PMS based LSP dataplane monitoring

Figure 1

For the sake of simplicity, let's assume that all the nodes are configured with the same SRGB [[I-D.ietf-spring-segment-routing](#)].

Let's assign the following Node SIDs to the nodes of the figure: PMS = 10, LER i = 20, LER j = 30.

To be able to work with the smallest possible SR label stack, first a suitable MPLS OAM method is used to detect the ECMP routed path between LER i to LER j which is to be monitored (and the required address information to direct a packet along it). Afterwards the PMS sets up and sends packets to monitor availability of the detected path. The PMS does this by creating a measurement packet with the following label stack (top to bottom): 20 - 30 - 10. The packet will only reliably use the monitored path, if the label and address information used in combination with the MPLS OAM method of choice is identical to that of the monitoring packet.

LER m forwards the packet received from the PMS to LSR1. Assuming Pen-ultimate Hop Popping to be deployed, LSR1 pops the top label and forwards the packet to LER i. There the top label has a value 30 and LER i forwards it to LER j. This will be done transmitting the packet via LSR1 or LSR2. The LSR will again pop the top label. LER j will forward the packet now carrying the top label 10 to the PMS (and it will pass a LSR and LER m).

A few observations on the example given in figure 1:

- o The path PMS to LER i must be available. This path must be detectable, but it is usually sufficient to apply a Shortest Path First algorithm based path.
- o If ECMP is deployed, it may be desired to measure along both possible paths which a packet may use between LER i and LER j. To do so, the MPLS OAM mechanism chosen to detect ECMP must reveal the required information (an example is a so called tree trace) between LER i and LER j. This method of dealing with ECMP based load balancing paths requires the smallest SR label stacks if monitoring of paths is applied after the tree trace completion.
- o The path LER j to PMS must be available. This path must be detectable, but it is usually sufficient to apply an SPF based path.

Once the MPLS paths (Node SIDs) and the required information to deal with ECMP has been detected, the paths of LER i to LER j can be monitored by the PMS. Monitoring itself does not require MPLS OAM functionality. All monitoring packets stay on dataplane, hence path monitoring does no longer require control plane interaction in any LER or LSR of the domain. To ensure reliable results, the PMS should be aware of any changes in IGP or MPLS topology. Further changes in ECMP functionality at LER i will impact results. Either the PMS should be notified of such changes or they should be limited to planned maintenance. After a topology change, a suitable MPLS OAM mechanism may be useful to detect the impact of the change.

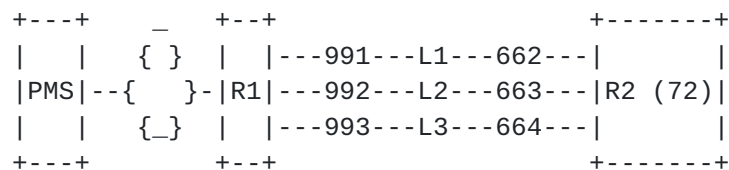
Determining a path to be executed prior to a measurement may also be done by setting up a label stack including all Node SIDs along that path (if LSR1 has Node SID 40 in the example and it should be passed between LER i and LER j, the label stack is 20 - 40 - 30 - 10). The advantage of this method is, that it does not involve MPLS OAM functionality and it is independent of ECMP functionalities. The method still is able to monitor all link combinations of all paths of an MPLS domain. If correct forwarding along the desired paths has to be checked, some suitable MPLS OAM mechanism may be applied also in this case.

In theory at least, a single PMS is able to monitor data plane availability of all LSPs in the domain. The PMS may be a router, but could also be dedicated monitoring system. If measurement system reliability is an issue, more than a single PMS may be connected to the MPLS domain.

Monitoring an MPLS domain by a PMS based on SR offers the option of monitoring complete MPLS domains with little effort and very excellent scalability. Data plane failure detection by circulating

monitoring packets can be executed at any time. The PMS further could be enabled to send MPLS OAM packets with the label stacks and address information identical to those of the monitoring packets to any node of the MPLS domain. Prior to monitoring a path, MPLS OAM may be used to detect ECMP dependant forwarding of a packet. A PMS may be designed to learn the IP address information required to execute a particular ECMP routed path and interfaces along that path. This allows to monitor these paths with label stacks reduced to a limited number of Node-SIDs resulting from SPF routing. The PMS does not require access to LSR/LER management interfaces or their control plane to do so.

4.2. Use Case 2 - Monitoring a Remote Bundle



SR based probing of all the links of a remote bundle

Figure 2

R1 addresses Lx by the Adjacency SID 99x, while R2 addresses Lx by the Adjacency SID 66(x+1).

In the above figure, the PMS needs to assess the dataplane availability of all the links within a remote bundle connected to routers R1 and R2.

The monitoring system retrieves the SID/Label information from the IGP LSDB and appends the following segment list/label stack: {72, 662, 992, 664} on its IP probe (whose source and destination addresses are the address of the PMS).

PMS sends the probe to its connected router. If the connected router is not SR compliant, a tunneling technique can be used to tunnel the probe and its MPLS stack to the first SR router. The MPLS/SR domain then forwards the probe to R2 (72 is the Node SID of R2). R2 forwards the probe to R1 over link L1 (Adjacency SID 662). R1 forwards the probe to R2 over link L2 (Adjacency SID 992). R2 forwards the probe to R1 over link L3 (Adjacency SID 664). R1 then forwards the IP probe to PMS as per classic IP forwarding.

4.3. Use Case 3 - Fault Localization

In the previous example, a uni-directional fault on the middle link in direction of R2 to R1 would be localized by sending the following two probes with respective segment lists:

- o 72, 662, 992, 664
- o 72, 663, 992, 664

The first probe would fail while the second would succeed. Correlation of the measurements reveals that the only difference is using the Adjacency SID 662 of the middle link from R1 to R2 in the non successful measurement. Assuming the second probe has been routed correctly, the fault must have been occurring in R2 which didn't forward the packet to the interface identified by its Adjacency SID 662.

5. Failure Notification from PMS to LERi

PMS on detecting any failure in the path liveliness may use any out-of-band mechanism to signal the failure to LER i. This document does not propose any specific mechanism and operators can choose any existing or new approach.

Alternately, the Operator may log the failure in local monitoring system and take necessary action by manual intervention.

6. Applying SR to Monitoring non-SR based LSPs (LDP and possibly RSVP-TE)

A SR based PMS connected to a MPLS domain consisting of LER and LSR supporting SR and LDP or RSVP-TE in parallel in all nodes may use SR paths to transmit packets to and from start and end points of non-SR based LSP paths to be monitored. In the above example, the label stack top to bottom may be as follows, when sent by the PMS:

- o Top: SR based Node-SID of LER i at LER m.
- o Next: LDP or RSVP-TE label identifying the path to LER j at LER i.
- o Bottom: SR based Node-SID identifying the path to the PMS at LER j

While the mixed operation shown here still requires the PMS to be aware of the LER LDP-MPLS or RSVP-TE topology, the PMS may learn the SR MPLS topology by IGP and use this information.

An implementation report on a PMS operating in an LDP domain is given in [[I-D.leipnitz-spring-pms-implementation-report](#)].

7. PMS Monitoring of Different Segment ID Types

MPLS SR topology awareness should allow the SID to monitor liveness of most types of SIDs (this may not be recommendable if a SID identifies an inter domain interface).

To match control plane information with data plane information, MPLS OAM functions as defined for example by [RFC 4379](#) [[RFC4379](#)] should be enhanced to allow collection of data relevant to check all relevant types of Segment IDs.

8. Connectivity Verification Using PMS

While the PMS based use cases explained in [Section 3](#) are sufficient to provide continuity check between LER i and LER j, it may not help perform connectivity verification. So in some cases like data plane programming corruption, it is possible that a transit node between LER i and LER j erroneously removes the top segment ID and forwards a monitoring packet to the PMS based on the bottom segment ID leading to a falsified path liveness indication by the PMS.

There are various method to perform basic connectivity verification like intermittently setting the TTL to 1 in bottom label so LER j selectively perform connectivity verification. Other methods are possible and may be added when requirements and solutions are specified.

9. Extensions of Specifications Relevant to this Use Case

The following activities are welcome enhancements supporting this use case, but they are not part of it:

[RFC 4379](#) [[RFC4379](#)] functions should be extended to support Flow- and Entropy Label based ECMP.

10. IANA Considerations

This memo includes no request to IANA.

11. Security Considerations

As mentioned in the introduction, a PMS monitoring packet should never leave the domain where it originated. It therefore should never use stale MPLS or IGP routing information. Further, assigning different label ranges for different purposes may be useful. A well

known global service level range may be excluded for utilisation within PMS measurement packets. These ideas shouldn't start a discussion. They rather should point out, that such a discussion is required when SR based OAM mechanisms like a SR are standardised.

12. Acknowledgements

The authors would like to thank Nobo Akiya for his contribution. Raik Leipnitz kindly provided an editorial review. The authors would also like to thank Faisal Iqbal for an insightful review and a useful set of comments and suggestions.

13. References

13.1. Normative References

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", [RFC 4379](https://www.rfc-editor.org/rfc/rfc4379), DOI 10.17487/RFC4379, February 2006, <<http://www.rfc-editor.org/info/rfc4379>>.

13.2. Informative References

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and j. jeffrant@gmail.com, "IS-IS Extensions for Segment Routing", [draft-ietf-isis-segment-routing-extensions-08](#) (work in progress), October 2016.

[I-D.ietf-mpls-spring-lsp-ping]
Kumar, N., Swallow, G., Pignataro, C., Akiya, N., Kini, S., Gredler, H., and M. Chen, "Label Switched Path (LSP) Ping/Trace for Segment Routing Networks Using MPLS Dataplane", [draft-ietf-mpls-spring-lsp-ping-00](#) (work in progress), May 2016.

[I-D.ietf-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", [draft-ietf-ospf-segment-routing-extensions-09](#) (work in progress), July 2016.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [draft-ietf-spring-segment-routing-09](#) (work in progress), July 2016.

[I-D.ietf-spring-sr-oam-requirement]

Kumar, N., Pignataro, C., Akiya, N., Geib, R., Mirsky, G., and S. Litkowski, "OAM Requirements for Segment Routing Network", [draft-ietf-spring-sr-oam-requirement-02](#) (work in progress), July 2016.

[I-D.leipzig-spring-pms-implementation-report]

Leipzig, R. and R. Geib, "A scalable and topology aware MPLS data plane monitoring system", [draft-leipzig-spring-pms-implementation-report-00](#) (work in progress), June 2016.

Authors' Addresses

Ruediger Geib (editor)
Deutsche Telekom
Heinrich Hertz Str. 3-7
Darmstadt 64295
Germany

Phone: +49 6151 5812747
Email: Ruediger.Geib@telekom.de

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium

Email: cfilsfil@cisco.com

Carlos Pignataro (editor)
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709-4987
US

Email: cpignata@cisco.com

Nagendra Kumar
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709
US

Email: naikumar@cisco.com

