INTERNET-DRAFT
Intended Status: Informational

Samer Salam
Ali Sajassi
Cisco
Sam Aldrin
Google
John E. Drake
Juniper
Donald Eastlake
Huawei

Expires: April 21, 2018 October 22, 2018

EVPN Operations, Administration and Maintenance Requirements and Framework
draft-salam-bess-evpn-oam-req-frmwk-01

Abstract

This document specifies the requirements and reference framework for Ethernet VPN (EVPN) Operations, Administration and Maintenance (OAM). The requirements cover the OAM aspects of EVPN and PBB-EVPN. The framework defines the layered OAM model encompassing the EVPN service layer, network layer and underlying Packet Switched Network (PSN) transport layer.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html. The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Copyright and License Notice

Table of Contents

1. Introduction

This document specifies the requirements and defines a reference framework for Ethernet VPN (EVPN) Operations, Administration and Maintenance (OAM, [RFC6291]). In this context, we use the term EVPN OAM to loosely refer to the OAM functions required for and/or applicable to [RFC7432] and [RFC7623].

EVPN is an Layer 2 VPN (L2VPN) solution for multipoint Ethernet services, with advanced multi-homing capabilities, using BGP for distributing customer/client MAC address reachability information over the core MPLS/IP network.

PBB-EVPN combines Provider Backbone Bridging (PBB) [802.1Q] with EVPN in order to reduce the number of BGP MAC advertisement routes, provide client MAC address mobility using C-MAC aggregation and B-MAC sub-netting, confine the scope of C-MAC learning to only active flows, offer per site policies, and avoid C-MAC address flushing on topology changes.

This document focuses on the fault management and performance management aspects of EVPN OAM.

1.1 Relationship to Other OAM Work

This document leverages concepts and draws upon elements defined and/or used in the following documents:

[RFC6136] specifies the requirements and a reference model for OAM as it relates to L2VPN services, pseudowires and associated Packet Switched Network (PSN) tunnels. This document focuses on VPLS and VPWS solutions and services.

[RFC8029] defines mechanisms for detecting data plane failures in MPLS LSPs, including procedures to check the correct operation of the data plane, as well as mechanisms to verify the data plane against the control plane.

[802.1Q] specifies the Ethernet Connectivity Fault Management (CFM) protocol, which defines the concepts of Maintenance Domains, Maintenance Associations, Maintenance End Points, and Maintenance Intermediate Points.

[Y.1731] extends Connectivity Fault Management in the following areas: it defines fault notification and alarm suppression functions for Ethernet. It also specifies mechanisms for Ethernet performance management, including loss, delay, jitter, and throughput measurement.

1.2 Specification of Requirements

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

1.3 Terminology

This document uses the following terminology defined in [RFC6136]:

MA Maintenance Association is a set of MEPs belonging to the same Maintenance Domain, established to verify the integrity of a single service instance.

MEP Maintenance End Point is responsible for origination and termination of OAM frames for a given MA.

MIP Maintenance Intermediate Point is located between peer MEPs and can process and respond to certain OAM frames but does not initiate them.

MD Maintenance Domain, an OAM Domain that represents a region over which OAM frames can operate unobstructed.

2. EVPN OAM Framework

2.1 OAM Layering

Multiple layers come into play for implementing an L2VPN service using the EVPN family of solutions:

- The Service Layer runs end to end between the sites or Ethernet Segments that are being interconnected by the EVPN solution.

- The Network Layer extends in between the EVPN PE nodes and is mostly transparent to the core nodes (except where Flow Entropy comes into play). It leverages MPLS for service (i.e. EVI) multiplexing and Split-Horizon functions.

- The Transport Layer is dictated by the networking technology of the PSN. It may be either based on MPLS LSPs or IP.

- The Link Layer is dependent upon the physical technology used. Ethernet is a popular choice for this layer, but other alternatives are deployed (e.g. POS, DWDM etc.).

This layering extends to the set of OAM protocols that are involved in the ongoing maintenance and diagnostics of EVPN networks. The figure below depicts the OAM layering, and shows which devices have visibility into what OAM layer(s).

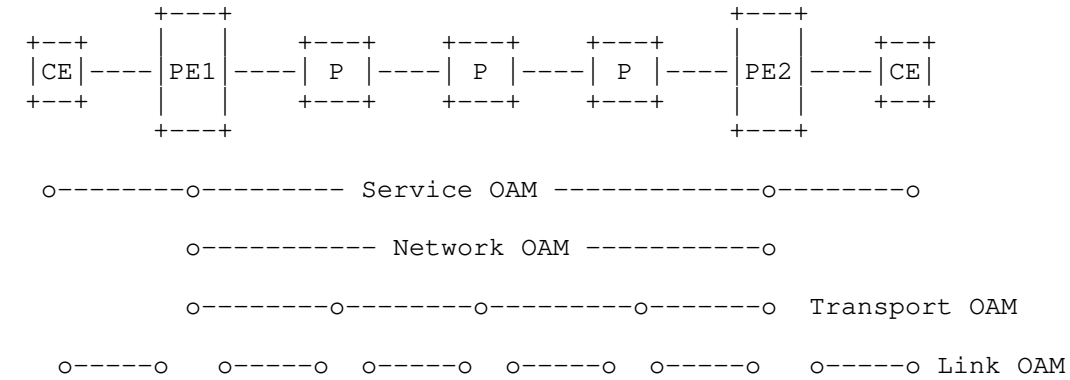

Figure 1: OAM Layering

Figure 2 below shows an example network where native Ethernet domains are interconnected via EVPN, and the OAM mechanisms applicable at each layer. The details of the layers are described in the sections below.

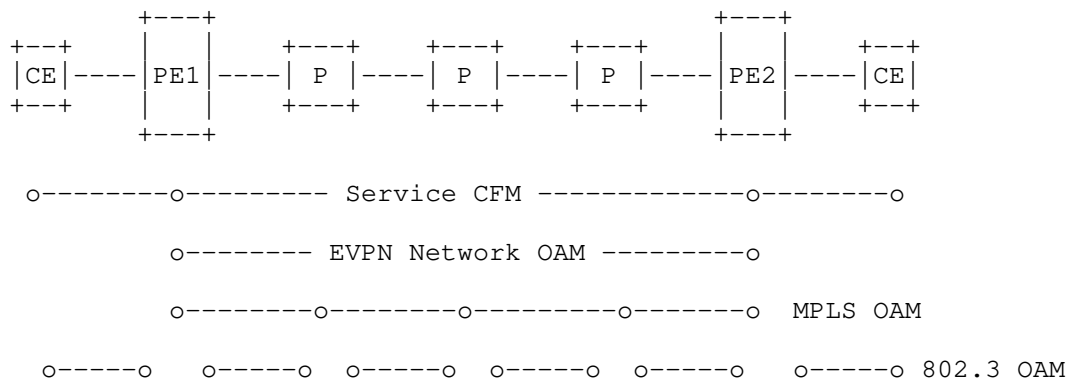

Figure 2: EVPN OAM Example

2.2 EVPN Service OAM

The EVPN Service OAM protocol depends on what service layer technology is being interconnected by the EVPN solution. In case of [RFC7432] and [RFC7623], the service layer is Ethernet; hence, the corresponding service OAM protocol is Ethernet Connectivity Fault Management (CFM) [802.1Q].

EVPN service OAM is visible to the CEs and EVPN PEs, but not to the core (P) nodes. This is because the PEs operate at the Ethernet MAC layer in [RFC7432] [RFC7623] whereas the P nodes do not.

The EVPN PE MUST support MIP functions in the applicable service OAM protocol, for example Ethernet CFM. The EVPN PE SHOULD support MEP functions in the applicable service OAM protocol. This includes both Up and Down MEP functions.

The EVPN PE MUST learn the MAC address of locally attached CE MEPs by snooping on CFM frames and advertising them to remote PEs as a MAC/IP Advertisement route.

The EVPN PE SHOULD advertise any MEP/MIP local to the PE as a MAC/IP Advertisement route. Since these are not subject to mobility, they SHOULD be advertised with the stick bit set (see Section 15.2 of [RFC7432]).

2.3 EVPN Network OAM

EVPN Network OAM is visible to the PE nodes only. This OAM layer is analogous to VCCV [RFC5085] in the case of VPLS/VPWS. It provides

mechanisms to check the correct operation of the data plane, as well as a mechanism to verify the data plane against the control plane. This includes the ability to perform fault detection and diagnostics on:

- the MP2P tunnels used for the transport of unicast traffic between PEs. EVPN allows for three different models of unicast label assignment: label per EVI, label per <ESI, Ethernet Tag> and label per MAC address. In all three models, the label is bound to an EVPN Unicast FEC.

EVPN Network OAM MUST provide mechanisms to check the operation of the data plane and verify that operation against the control plane view.

- the MP2P tunnels used for aliasing unicast traffic destined to a multi-homed Ethernet Segment. The three label assignment models, discussed above, apply here as well. In all three models, the label is bound to an EVPN Aliasing FEC. EVPN Network OAM MUST provide mechanisms to check the operation of the data plane and verify that operation against the control plane view.

- the multicast tunnels (either MP2P or P2MP) used for the transport of broadcast, unknown unicast and multicast traffic between PEs. In the case of ingress replication, a label is allocated per EVI or per <EVI, Ethernet Tag> and is bound to an EVPN Multicast FEC. In the case of LSM, and more specifically aggregate inclusive trees, again a label may be allocated per EVI or per <EVI, Ethernet Tag> and is bound to the tunnel FEC.

- the correct operation of the ESI split-horizon filtering function. In EVPN, a label is allocated per multi-homed Ethernet Segment for the purpose of performing the access split-horizon enforcement. The label is bound to an EVPN Ethernet Segment.

- the correct operation of the DF filtering function.

EVPN Network OAM MUST provide mechanisms to check the operation of the data plane and verify that operation against the control plane view for the DF filtering function.

EVPN network OAM mechanisms MUST provide in-band management capabilities. As such, OAM messages MUST be encoded so that they exhibit identical entropy characteristics to data traffic.

EVPN network OAM SHOULD provide both proactive and on-demand mechanisms of monitoring the data plane operation and data plane conformance to the state of the control plane.

2.4 Transport OAM for EVPN

The transport OAM protocol depends on the nature of the underlying transport technology in the PSN. MPLS OAM mechanisms [RFC8029] [RFC6425] as well as ICMP [RFC792] are applicable, depending on whether the PSN employs MPLS or IP transport, respectively. Furthermore, BFD mechanisms per [RFC5880], [RFC5881], [RFC5883] and [RFC5884] apply. Also, the BFD mechanisms pertaining to MPLS-TP LSPs per [RFC6428] are applicable.

2.5 Link OAM

Link OAM depends on the data link technology being used between the PE and P nodes. For example, if Ethernet links are employed, then Ethernet Link OAM [802.3] Clause 57 may be used.

2.6 OAM Inter-working

When inter-working two networking domains, such as native Ethernet and EVPN to provide an end-to-end emulated service, there is a need to identify the failure domain and location, even when a PE supports both the Service OAM mechanisms and the EVPN Network OAM mechanisms. In addition, scalability constraints may not allow running proactive monitoring, such as Ethernet Continuity Check Messages (CCMs), at a PE to detect the failure of an EVI across the EVPN domain. Thus, the mapping of alarms generated upon failure detection in one domain (e.g. native Ethernet or EVPN network domain) to the other domain is needed. There are also cases where a PE may not be able to process Service OAM messages received from a remote PE over the PSN even when such messages are defined, as in the Ethernet case, thereby necessitating support for fault notification message mapping between the EVPN Network domain and the Service domain.

OAM inter-working is not limited though to scenarios involving disparate network domains. It is possible to perform OAM inter-working across different layers in the same network domain. In general, alarms generated within an OAM layer, as a result of proactive fault detection mechanisms, may be injected into its client layer OAM mechanisms. This allows the client layer OAM to trigger event-driven (i.e. asynchronous) fault notifications. For example, alarms generated by the Link OAM mechanisms may be injected into the Transport OAM layer, and alarms generated by the Transport OAM mechanism may be injected into the Network OAM mechanism, and so on.

EVPN OAM MUST support inter-working between the Network OAM and Service OAM mechanisms. EVPN OAM MAY support inter-working among

other OAM layers.

3. EVPN OAM Requirements

This section discusses the EVPN OAM requirements pertaining to Fault Management and Performance Management.

3.1 Fault Management Requirements

3.1.1 Proactive Fault Management Functions

The network operator configures proactive fault management functions to run periodically without a time bound. Certain actions, for example protection switchover or alarm indication signaling, can be associated with specific events, such as entering or clearing fault states.

3.1.1.1 Fault Detection (Continuity Check)

Proactive fault detection is performed by periodically monitoring the reachability between service endpoints, i.e. MEPs in a given MA, through the exchange of Continuity Check messages. The reachability between any two arbitrary MEPs may be monitored for:

- in-band per-flow monitoring. This enables per flow monitoring between MEPs. EVPN Network OAM MUST support fault detection with per user flow granularity. EVPN Service OAM MAY support fault detection with per user flow granularity.

- a representative path. This enables liveness check of the nodes hosting the MEPs assuming that the loss of continuity to the MEP is interpreted as a failure of the hosting node. This, however, does not conclusively indicate liveness of the path(s) taken by user data traffic. This enables node failure detection but not path failure detection, through the use of a test flow. EVPN Network OAM and Service OAM MUST support fault detection using test flows.

- all paths. For MPLS/IP networks with ECMP, monitoring of all unicast paths between MEPs (on non-adjacent nodes) may not be possible, since the per-hop ECMP hashing behavior may yield situations where it is impossible for a MEP to pick flow entropy characteristics that result in exercising the exhaustive set of ECMP paths. Monitoring of all ECMP paths between MEPs (on non-adjacent nodes) is not a requirement for EVPN OAM.

The fact that MPLS/IP networks do not enforce congruency between

unicast and multicast paths means that the proactive fault detection mechanisms for EVPN networks MUST provide procedures to monitor the unicast paths independently of the multicast paths. This applies to EVPN Service OAM and Network OAM.

3.1.1.2 Defect Indication

EVPN Service OAM MUST support event-driven defect indication upon the detection of a connectivity defect. Defect indications can be categorized into two types: forward and reverse defect indications.

3.1.1.2.1 Forward Defect Indication

This is used to signal a failure that is detected by a lower layer OAM mechanism. A server MEP (i.e. an actual or virtual MEP) transmits a Forward Defect Indication in a direction that is away from the direction of the failure (refer to Figure 3 below).

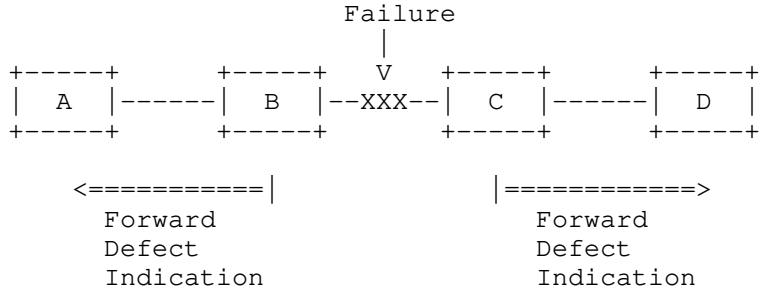

Figure 3: Forward Defect Indication

Forward defect indication may be used for alarm suppression and/or for purpose of inter-working with other layer OAM protocols. Alarm suppression is useful when a transport/network level fault translates to multiple service or flow level faults. In such a scenario, it is enough to alert a network management station (NMS) of the single transport/network level fault in lieu of flooding that NMS with a multitude of Service or Flow granularity alarms. EVPN PEs SHOULD support Forward Defect Indication in the Service OAM mechanisms.

3.1.1.2.2 Reverse Defect Indication (RDI)

RDI is used to signal that the advertising MEP has detected a loss of continuity (LoC) defect. RDI is transmitted in the direction of the

failure (refer to Figure 4).

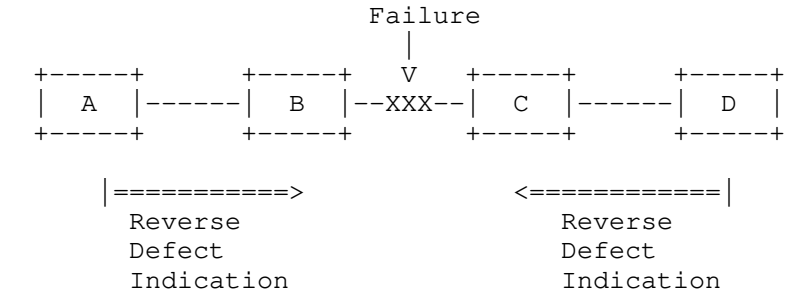

Figure 4: Reverse Defect Indication

RDI allows single-sided management, where the network operator can examine the state of a single MEP and deduce the overall health of a monitored service. EVPN PEs SHOULD support Reverse Defect Indication in the Service OAM mechanisms. This includes both the ability to signal LoC defect to a remote MEP, as well as the ability to recognize RDI from a remote MEP. Note that, in a multipoint MA, RDI is not a useful indicator of unidirectional fault. This is because RDI carries no indication of the affected MEP(s) with which the sender had detected a LoC defect.

3.1.2 On-Demand Fault Management Functions

On-demand fault management functions are initiated manually by the network operator and continue for a time bound period. These functions enable the operator to run diagnostics to investigate a defect condition.

3.1.2.1 Connectivity Verification

EVPN Network OAM MUST support on-demand connectivity verification mechanisms for unicast and multicast destinations. The connectivity verification mechanisms SHOULD provide a means for specifying and carrying in the messages:

- variable length payload/padding to test MTU related connectivity problems.

- test frame formats as defined in Appendix C of [RFC2544] to detect potential packet corruption.

EVPN Network OAM MUST support connectivity verification at per flow

granularity. This includes both user flows (to test a specific path between PEs) as well as test flows (to rest a representative path between PEs).

EVPN Service OAM MUST support connectivity verification on test flows and MAY support connectivity verification on user flows.

For multicast connectivity verification, EVPN Network OAM MUST support reporting on:

- the DF filtering status of specific port(s) or all the ports in a given bridge-domain.

- the Split Horizon filtering status of specific port(s) or all the ports in a given bridge-domain.

3.1.2.2 Fault Isolation

EVPN OAM MUST support an on-demand fault localization function. This involves the capability to narrow down the locality of a fault to a particular port, link or node. The characteristic of forward/reverse path asymmetry, in MPLS/IP, renders fault isolation into a direction-sensitive operation. That is, given two PEs A and B, localization of continuity failures between them requires running fault isolation procedures from PE A to PE B as well as from PE B to PE A.

EVPN Service OAM mechanisms only have visibility to the PEs but not the MPLS/IP P nodes. As such, they can be used to deduce whether the fault is in the customer's own network, the local CE-PE segment or remote CE-PE segment(s). EVPN Network and Transport OAM mechanisms can be used for fault isolation between the PEs and P nodes.

3.2 Performance Management

Performance Management functions can be performed both proactively and on-demand. Proactive management involves a recurring function, where the performance management probes are run continuously without a trigger. We cover both proactive and on-demand functions in this section.

3.2.1 Packet Loss

EVPN Network OAM SHOULD provide mechanisms for measuring packet loss for a given service.

Given that EVPN provides inherent support for multipoint-to-multipoint connectivity, then packet loss cannot be accurately measured by means of counting user data packets. This is because user packets can be delivered to more PEs or more ports than are necessary (e.g. due to broadcast, un-pruned multicast or unknown unicast flooding). As such, a statistical means of approximating packet loss rate is required. This can be achieved by sending "synthetic" OAM packets that are counted only by those ports (MEPs) that are required to receive them. This provides a statistical approximation of the number of data frames lost, even with multipoint-to-multipoint connectivity.

3.2.2 Packet Delay

EVPN Service OAM SHOULD support measurement of one-way and two-way packet delay and delay variation (jitter) across the EVPN network. Measurement of one-way delay requires clock synchronization between the probe source and target devices. Mechanisms for clock synchronization are outside the scope of this document. Note that Service OAM performance management mechanisms defined in [Y.1731] can be used.

EVPN Network OAM MAY support measurement of one-way and two-way packet delay and delay variation (jitter) across the EVPN network.

4. Security Considerations

EVPN OAM must provide mechanisms for:

- Preventing denial of service attacks caused by exploitation of the OAM message channel.

- Optionally authenticate communicating endpoints (MEPs and MIPs)

- Preventing OAM packets from leaking outside of the EVPN network or outside their corresponding Maintenance Domain. This can be done by having MEPs implement a filtering function based on the Maintenance Level associated with received OAM packets.

5. Acknowledgements

The authors would like to thank the following for their review of this work and valuable comments:

Gregory Mirsky, Alexander Vainshtein

6. IANA Considerations

This document requires no IANA actions.

Normative References

[RFC792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <https://www.rfc-editor.org/info/rfc792>.

[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>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

[RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881, DOI 10.17487/RFC5881, June 2010, <https://www.rfc-editor.org/info/rfc5881>.

[RFC5883] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for Multihop Paths", RFC 5883, DOI 10.17487/RFC5883, June 2010, <https://www.rfc-editor.org/info/rfc5883>.

[RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, DOI 10.17487/RFC5884, June 2010, <https://www.rfc-editor.org/info/rfc5884>.<

[RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu, D., and S. Mansfield, "Guidelines for the Use of the "OAM" Acronym in the IETF", BCP 161, RFC 6291, DOI 10.17487/RFC6291, June 2011, <https://www.rfc-editor.org/info/rfc6291>.

[RFC6425] Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011, <https://www.rfc-editor.org/info/rfc6425>.

[RFC6428] Allan, D., Ed., Swallow, G., Ed., and J. Drake, Ed., "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428, DOI 10.17487/RFC6428, November 2011, <https://www.rfc-editor.org/info/rfc6428>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February

2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC7623] Sajassi, A., Ed., Salam, S., Bitar, N., Isaac, A., and W. Henderickx, "Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)", RFC 7623, DOI 10.17487/RFC7623, September 2015, <https://www.rfc-editor.org/info/rfc7623>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <http://www.rfc-editor.org/info/rfc8174>

Informative References

[802.1Q] "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Bridges and Virtual Bridge Local Area Networks", 2014.

[Y.1731] "ITU-T Recommendation Y.1731 (02/08) - OAM functions and mechanisms for Ethernet based networks", February 2008.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, DOI 10.17487/RFC2544, March 1999, <https://www.rfc-editor.org/info/rfc2544>.

[RFC5085] Nadeau, T., Ed., and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, DOI 10.17487/RFC5085, December 2007, <https://www.rfc-editor.org/info/rfc5085>.

[RFC6136] Sajassi, A., Ed., and D. Mohan, Ed., "Layer 2 Virtual Private Network (L2VPN) Operations, Administration, and Maintenance (OAM) Requirements and Framework", RFC 6136, DOI 10.17487/RFC6136, March 2011, <https://www.rfc-editor.org/info/rfc6136>.

Authors' Addresses

Samer Salam
Cisco

Email: ssalam@cisco.com

Ali Sajassi
Cisco
170 West Tasman Drive
San Jose, CA 95134, USA

Email: sajassi@cisco.com

Sam Aldrin
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA USA

Email: aldrin.ietf@gmail.com

John E. Drake
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

Email: jdrake@juniper.net

Donald E. Eastlake, 3rd
Huawei Technologies
1424 Pro Shop Court
Davenport, FL 33896 USA

Tel: +1-508-333-2270
Email: d3e3e3@gmail.com