

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: November 7, 2020

S. Jacob, Ed.
K. Tiruveedhula
Juniper Networks
May 6, 2020

Benchmarking Methodology for EVPN VPWS
draft-kishjac-bmwg-evpnvpwstest-04

Abstract

This document defines methodologies for benchmarking EVPN-VPWS performance. EVPN-VPWS is defined in RFC 8214, and is being deployed in Service Provider networks. Specifically this document defines the methodologies for benchmarking EVPN-VPWS Scale convergence, Fail over, Core isolation, high availability and longevity.

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 <https://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 November 7, 2020.

Copyright Notice

Copyright (c) 2020 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 (<https://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. Introduction	2
1.1. Requirements Language	2
1.2. Terminologies	2
2. Test Topology	4
3. Test Cases	6
3.1. Local Failure Scenario 1	7
3.2. Local Failure Scenario 2	7
3.3. Core Failure	8
3.4. Link Flap	9
4. Scale Convergence	9
4.1. To measure the packet loss during the core link failure.	9
5. High Availability	10
5.1. To Record the whether there is traffic loss due to routing engine failover for redundancy test.	10
6. SOAK Test	11
6.1. To Measure the stability of the DUT with scale and traffic.	11
7. Acknowledgements	12
8. IANA Considerations	12
9. Security Considerations	12
10. References	12
10.1. Normative References	12
10.2. Informative References	12
Appendix A. Appendix	13
Authors' Addresses	13

1. Introduction

EVPN-VPWS is defined in RFC 8214, discusses how VPWS can be combined with EVPNs to provide a new/combined solution. This draft defines methodologies that can be used to benchmark RFC 8214 solutions. Further, this draft provides methodologies for benchmarking the performance of EVPN VPWS Scale, Scale Convergence, Core isolation, longevity, high availability.

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].

1.2. Terminologies

All-Active Redundancy Mode: When all PEs attached to an Ethernet segment are allowed to forward known unicast traffic to/from that

Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in All-Active redundancy mode.

AA: All Active mode

AC: Attachment Circuits

CE: Customer Router/Devices/Switch.

DF: Designated Forwarder

DUT: Device under test.

Ethernet Segment (ES): When a customer site (device or network) is connected to one or more PEs via a set of Ethernet links, then that set of links is referred to as an 'Ethernet segment'.

EVI: An EVPN instance spanning the Provider Edge (PE) devices participating in that EVPN.

Ethernet Segment Identifier (ESI): A unique non-zero identifier that identifies an Ethernet segment is called an 'Ethernet Segment Identifier'.

Ethernet Tag: An Ethernet tag identifies a particular broadcast domain, e.g., a VLAN. An EVPN instance consists of one or more broadcast domains.

Interface: Physical interface of a router/switch.

IRB: Integrated routing and bridging interface

MAC: Media Access Control addresses on a PE.

MHPE2: Multi homed Provider Edge router 2.

MHPE1: Multi homed Provider Edge router 1.

SHPE3: Single homed Provider Edge Router 3.

PE: Provider Edge device.

P: Provider Router.

RR: Route Reflector.

RT: Traffic Generator.

Sub Interface: Each physical Interfaces is subdivided into Logical units.

SA: Single Active

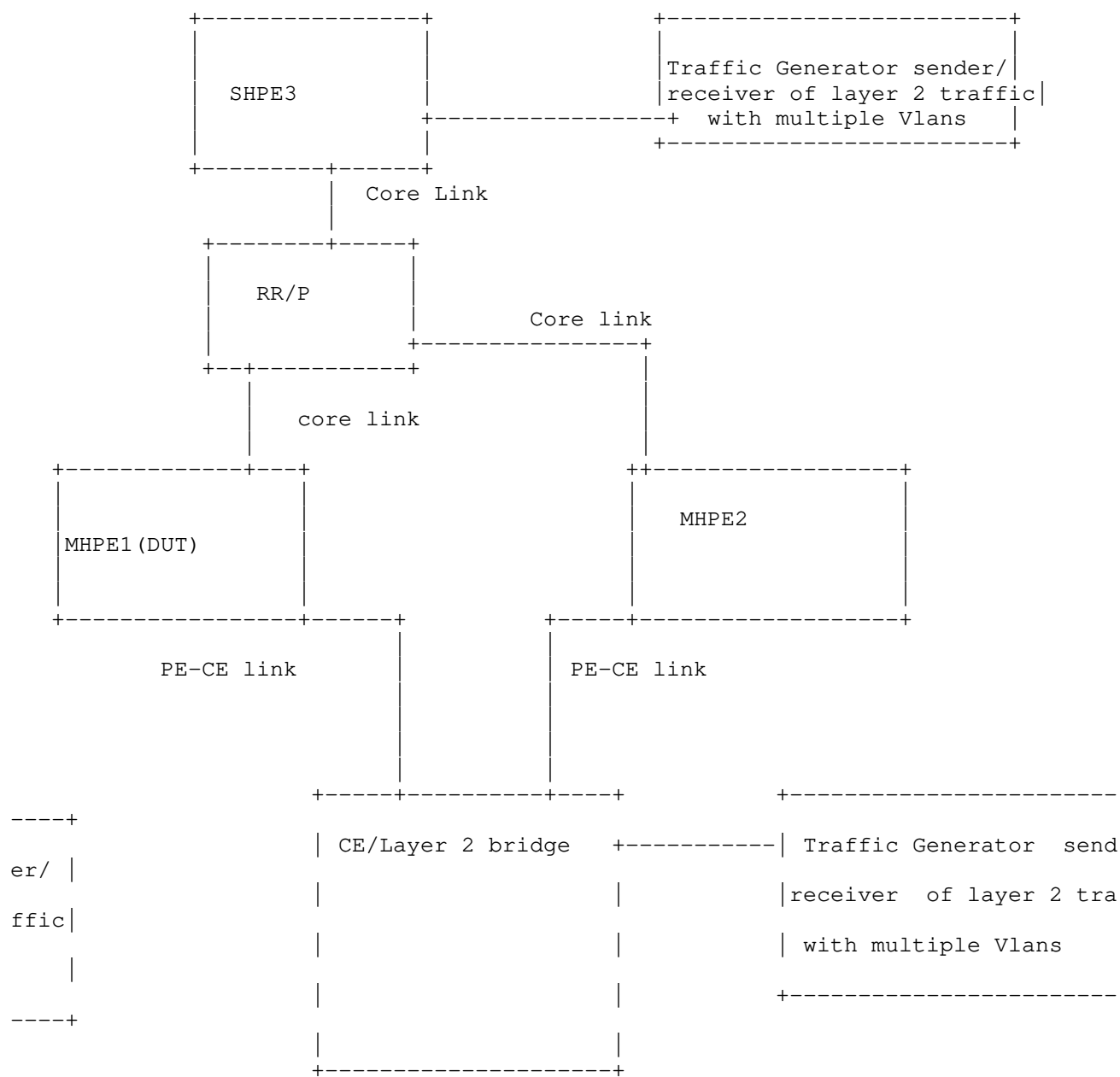
Single-Active Redundancy Mode: When only a single PE, among all the PEs attached to an Ethernet segment, is allowed to forward traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in Single-Active redundancy mode.

VPWS: Virtual private wire service.

2. Test Topology

There are five routers in the Test setup. SHPE3, RR/P, MHPE1 and MHPE2 emulating a service provider network. CE is a customer device connected to MHPE1 and MHPE2, it is configured with bridge domains in multiple vlans. The traffic generator is connected to CE and SHPE3. The MHPE1 acts as DUT. The traffic generator will be used as sender and receiver of traffic. The DUT will be the reference point for all the test cases. MHPE1 and MHPE2 are multi home routers connected to CE running single active mode. The traffic generator will be generating traffic at 10% of the line rate.

Topology Diagram



Topology 1

Topology Diagram

Figure 1

Test Setup Configurations:

SHPE3 is configured with Interior Gateway protocols like OPSF or IS-IS for underlay, LDP for MPLS support, Interior Border Gateway with EVPN address family for overlay support. This router must be configured with N EVPN-VPWS instances for testing. Traffic generator is connected to this router for sending and receiving traffic.

RR is configured with Interior Gateway protocols like OPSF or IS-IS for underlay, LDP for MPLS support, Interior Border Gateway with EVPN address family for overlay support. This router acts as a provider router and as a route reflector.

MHPE1 is configured with Interior Gateway protocols like OPSF or IS-IS for underlay, LDP for MPLS support, Interior Border Gateway with EVPN address family for overlay support. This router must be configured with N EVPN-VPWS instances for testing. This router is configured with ESI per vlan or ESI per interface. It is functioning as multi homing PE working on Single Active EVPN mode. This router serves as the DUT and it is connected to CE. MHPE1 is acting as DUT for all the test cases.

MHPE2 is configured with Interior Gateway protocols like OPSF or IS-IS for underlay, LDP for MPLS support, Interior Border Gateway with EVPN address family for overlay support. This router must be configured with N EVPN-VPWS instances for testing. This router is configured with ESI per vlan or ESI per interface. It is functioning as multi homing PE working on Single Active EVPN mode. It is connected to CE.

CE is acting as bridge configured with multiple vlans, the same vlans are configured on MHPE1, MHPE2, SHPE3. traffic generator is connected to CE. The traffic generator acts as sender or receiver of traffic.

Depending up on the test scenarios the traffic generators will be used to generate uni directional or bi directional flows.

The above configuration will be serving as the base configuration for all test cases.

3. Test Cases

The following tests are conducted to measure the packet loss during the local link and core failure in DUT with Scaled AC's.

3.1. Local Failure Scenario 1

Objective:

Measure the time taken to switch from primary to backup during local link failure.

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to CE from traffic generator. The traffic is uni directional and it flows from CE to DUT working as Active router. Then shut the DUT-CE link, so that traffic from CE switches to MHPE2. Traffic must be tested with various line rate that from 10% to 98%.

Measurement :

Measure the time taken by the traffic to switch from Active router to the backup. The test is repeated for "N" times and the values are collected. The AC's local switch over time is calculated by averaging the values obtained by "N" samples. "N" is an arbitrary number to get a sufficient sample. The time measured for each sample is denoted by T1, T2...Tn. The measurement is carried out using external server which polls the DUT using automated scripts. Fail over time must be measured for various line rate.

AC's switch over from primary to backup PE in sec = $(T1+T2+...Tn/N)$

3.2. Local Failure Scenario 2

Objective:

Measure time taken by remote PE to switch traffic from primary to backup during CE link failure.

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to SHPE3 from traffic generator. The traffic is uni directional and it flows from

SHPE3 to DUT working as Active router. Then shut the DUT-CE link, so the remote traffic flow switches from DUT to MHPE2. Traffic must be tested with various line rate that from 10% to 98%.

Measurement :

Measure the time taken by the traffic to switch from Active router to the backup. The test is repeated for "N" times and the values are collected. The AC's switch over time for the remote traffic is calculated by averaging the values obtained by "N" samples. "N" is an arbitrary number to get a sufficient sample. The time measured for each sample is denoted by T1,T2...Tn. The measurement is carried out using external server which polls the DUT using automated scripts. Fail over time must be measured for various line rate.

AC's switch over from primary to backup PE in sec = $(T1+T2+..Tn/N)$

3.3. Core Failure

Objective:

Measure the time taken by remote PE to switch traffic from primary to backup during core link failure.

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to SHPE3 from traffic generator. The traffic is uni directional and it flows from SHPE3 to DUT working as Active router. Then shut the DUT core link, so the remote traffic flow switches from DUT to MHPE2. Traffic must be tested with various line rate that from 10% to 98%.

Measurement :

Measure the time taken by the traffic to switch from Active router to the backup. The test is repeated for "N" times and the values are collected. The AC's switch over time for the remote traffic is calculated by averaging the values obtained by "N" samples. "N" is an arbitrary number to get a sufficient sample. The time measured for each sample is denoted by T1,T2...Tn. The measurement is carried out using external server which polls the DUT using automated scripts. Fail over time must be measured for various line rate.

AC's core Failure fail over time = $(T1+T2+...Tn/N)$

3.4. Link Flap

Objective:

Measure time taken by primary PE to regain control after the local PE-CE link flap.

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to CE from traffic generator. The traffic is uni directional and it flows from CE to DUT working as Active router. Then shut the DUT core link, so the local traffic flow switches from DUT to MHPE2. Once the fail over is performed. Bring the link up. Now the DUT becomes the Active router. Measure time taken by the DUT to regain the traffic. Traffic must be tested with various line rate that from 10% to 98%.

Measurement :

Measure the time taken by the traffic to switch back to the DUT. The test is repeated for "N" times and the values are collected. The AC's switch over time for the remote traffic is calculated by averaging the values obtained by "N" samples. "N" is an arbitrary number to get a sufficient sample. The time measured for each sample is denoted by T1, T2...Tn. The measurement is carried out using external server which polls the DUT using automated scripts. Fail over time must be measured for various line rate.

Time taken to switch back to primary (DUT) once the link is restored = $(T1+T2+...Tn/N)$

4. Scale Convergence

4.1. To measure the packet loss during the core link failure.

Objective:

Measure the convergence at a higher number of AC's

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N*100" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to CE from traffic generator and send traffic from traffic generator to SHPE3. The traffic is directional and it flows from CE to DUT and from DUT to CE, working as Active router. Then shut the DUT core link, so the traffic flow switches from DUT to MHPE2. Measure traffic switching time. Traffic must be tested with various line rate that from 10% to 98%.

Measurement :

Measure the time taken by the traffic to switch from DUT to MHPE2. The test is repeated for "N" times and the values are collected. The AC's switch over time for the traffic is calculated by averaging the values obtained by "N" samples. "N" is an arbitrary number to get a sufficient sample. The time measured for each sample is denoted by T1, T2...Tn. The measurement is carried out using external server which polls the DUT using automated scripts. Fail over time must be measured for various line rate.

Packet loss in sec = $(T1+T2+...Tn/N)$

5. High Availability

5.1. To Record the whether there is traffic loss due to routing engine failover for redundancy test.

Objective:

Measure the traffic loss during routing engine fail over.

Topology : Topology 1

Procedure:

Confirm the DUT is up and running with EVPN-VPWS. The AC must be up and running. "N*100" AC's in MHPE1, MHPE2, working in SA mode. Ensure DUT is active and MHPE2 is backup PE. Send unicast packets to CE and SHPE3 from traffic generator. The traffic is directional and it flows from CE to DUT and from DUT to CE, working as Active router. Do a routing engine fail over once the traffic is stabilized in DUT. Traffic must be tested with various line rate that from 10% to 98%. The expectation is 0 packet loss, no role change in AC's.

Measurement :

The expectation of the test is 0 traffic loss with no change in the DF role. DUT should not withdraw any routes. But in cases where the DUT is not properly synchronized between master and standby, due to that packet loss are observed. In that scenario the packet loss is measured. The test is repeated for "N" times and the values are collected. The packet loss is calculated by averaging the values obtained by "N" samples.

Packet loss in sec = $(T1+T2+..Tn/N)$

6. SOAK Test

This test is carried out to measure the stability of the DUT in a scaled environment with traffic over a period of time "T' ". In each interval "t1" the DUT CPU usage, memory usage are measured. The DUT is checked for any crashes during this time period.

6.1. To Measure the stability of the DUT with scale and traffic.

Objective:

To measure the stability of the DUT in a scaled environment with traffic.

Topology : Topology 1

Procedure:

Scale N AC's in DUT, SHPE3 and MHPE2. Send F frames to DUT from CE using traffic generator with different X SA and DA for N EVI's. Send F frames from traffic generator to SHPE3 with X different SA and DA. There is a bi directional traffic flow with F pps in each direction. The DUT must run with traffic for 24 hours, every hour check for memory leak, crash.

Measurement :

Take the hourly reading of CPU, process memory. There should not be any leak, crashes, CPU spikes. The CPU spike is determined as the CPU usage which shoots at 40 to 50 percent of the average usage. The average value vary from device to device. Memory leak is determined by increase usage of the memory for EVPN-VPWS process. The expectation is under steady state the memory usage for EVPN-VPWS process should not increase.

7. Acknowledgements

We would like to thank Al and Sarah for the support.

8. IANA Considerations

This memo includes no request to IANA.

9. Security Considerations

The benchmarking tests described in this document are limited to the performance characterization of controllers in a lab environment with isolated networks. The benchmarking network topology will be an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network or misroute traffic to the test management network. Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the controller. Special capabilities SHOULD NOT exist in the controller specifically for benchmarking purposes. Any implications for network security arising from the controller SHOULD be identical in the lab and in production networks.

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>>.
- [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>>.
- [RFC2899] Ginoza, S., "Request for Comments Summary RFC Numbers 2800-2899", RFC 2899, DOI 10.17487/RFC2899, May 2001, <<https://www.rfc-editor.org/info/rfc2899>>.

10.2. Informative References

- [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>>.

[RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <<https://www.rfc-editor.org/info/rfc8214>>.

Appendix A. Appendix

Authors' Addresses

Sudhin Jacob (editor)
Juniper Networks
Bangalore
India

Phone: +91 8061212543
Email: sjacob@juniper.net

Kishore Tiruveedhula
Juniper Networks
10 Technology Park Dr
Westford, MA 01886
USA

Phone: +1 9785898861
Email: kishoret@juniper.net